

HERIOT-WATT UNIVERSITY

# Effects of Genetic Testing on Insurance: Pedigree Analysis and Ascertainment Adjustment

Laura MacCalman

SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN ACTUARIAL SCIENCE

ON COMPLETION OF RESEARCH IN THE

DEPARTMENT OF ACTUARIAL MATHEMATICS AND STATISTICS,

SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES.

March 2009

The copyright in this thesis is owned by the author. Any quotation from the thesis or use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information.

# Declaration Statement

I hereby declare that the work presented in this thesis was carried out by myself at Heriot-Watt University, Edinburgh, except where due acknowledgement is made, and has not been submitted for any other degree.

---

Laura MacCalman (Candidate)

---

Professor Angus S Macdonald (Supervisor)

---

Date

# Abstract

Recent advances in genetics have resulted in the identification of mutations responsible for a number of genetic disorders which, in turn, have led to the development of genetic tests. The use of genetic testing raises the issue of who should be allowed access to the results; in particular should insurers be allowed to use genetic test results when calculating premium rates? At the moment there is a self-imposed moratorium in the UK preventing insurers from using the results of presymptomatic genetic tests until more investigation is carried out.

It is well-established that critical-illness (and sometimes life) insurance cannot be offered to mutation carriers. However such conclusions have (necessarily) been based on the published medical studies available, few of which include the detail needed to reconstruct the data. At the same time, the Genetics and Insurance Committee (GAIC) is setting out criteria that must be met if any genetic tests may be used in underwriting. These criteria cover questions of reliability that from a statistical point of view must include the estimation of insurance premiums from medical or epidemiological data. This question is rarely addressed: we address it in this thesis using pedigree data for Huntington's Disease and BRCA1-related breast and ovarian cancer.

In particular, we study the extent to which ascertainment bias, long known to affect pedigree analysis, affects the actuarial questions of pricing insurance. Having direct access to pedigree data gives us a unique opportunity to analyse how the sampling uncertainty inherent in the data translates into sampling uncertainty in actuarial quantities such as premium rates; moreover, allowing for ascertainment bias and adjustments to remove it. In particular, we are able to assess the validity of *ad hoc* adjustments to onset rates used by other authors.

# Acknowledgements

I would like to thank my supervisor Angus Macdonald for all of his guidance and support throughout the study. He always made time to meet with me and helped the quality of my work and this thesis to improve through many discussions and countless suggestions, without his help and encouragement this thesis may never have been completed.

I would also like to thank Jacki Needs and Professor Lindsay Prior who were involved in this investigation from an early stage. It was Jacki who searched through countless medical records and managed to put together the pedigrees, without which none of this work could have been carried out. Also thanks to the Wellcome Trust who funded this work initially and to the Actuarial mathematics and statistics department of Heriot-Watt University for the funding which enabled me to finish this work.

It has also been a pleasure to work with and share offices with my colleagues Alex Cook, Pradip Tapadar, Tunde Akodu and Lu Li. Thanks to them for their help, and company, throughout my years at Heriot-Watt University.

I would like to thank my mum and dad and the rest of my family, and ‘extended’ family, (far too many to name, but you know who you are) for all of their love and support, particularly throughout my many years studying.

Finally I would like to dedicate this work to my fiancé Frazer and our new family. Without Frazer’s continued support I would not have got this far and the imminent arrival of our first gorgeous baby, Chloe, provided the motivation to finish this work as I was (rightly) told that I would have little time (or energy) to do so after her arrival.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Genetics</b>	<b>5</b>
2.1 An Introduction to Genetics . . . . .	5
2.1.1 DNA . . . . .	5
2.1.2 Chromosomes . . . . .	9
2.2 Inheritance . . . . .	9
2.2.1 Cell Division and Reproduction . . . . .	9
2.2.2 Mendel's Laws of Inheritance . . . . .	10
2.2.3 Mutations . . . . .	11
2.2.4 Disorders . . . . .	12
2.3 Pedigrees . . . . .	14
2.3.1 Information That Can Be Obtained From Pedigrees . . . . .	15
2.3.2 Ascertainment Bias . . . . .	17
<b>3 Insurance</b>	<b>19</b>
3.1 Insurance and Risk . . . . .	19

3.2	Use of Genetic Information by UK Insurers . . . . .	20
3.3	Information Available About CI and Life Insurance Risks . . . . .	23
3.4	Adverse Selection . . . . .	25
3.5	Research Concerning Genetics and Insurance . . . . .	26
3.5.1	Breast and Ovarian Cancer . . . . .	26
3.5.2	Huntington's Disease . . . . .	32
3.5.3	General Discussion . . . . .	35
<b>4</b>	<b>Epidemiological Review</b>	<b>37</b>
4.1	Huntington's Disease . . . . .	37
4.1.1	Genetic Testing and Penetrance . . . . .	38
4.1.2	Onset Rates . . . . .	41
4.2	Breast and Ovarian Cancer . . . . .	44
4.2.1	Genetic Testing . . . . .	45
4.2.2	Prevalence . . . . .	46
4.2.3	Penetrance . . . . .	48
4.2.4	Treatment . . . . .	49
4.2.5	Other Genes . . . . .	51
4.3	Familial Adenomatous Polyposis . . . . .	51
4.3.1	Genetic Testing . . . . .	53
4.3.2	Diagnosis of FAP . . . . .	54
4.3.3	Colorectal Cancer . . . . .	55
4.3.4	Other Manifestations of FAP . . . . .	56
4.3.5	Treatment . . . . .	57
4.3.6	Survival . . . . .	58

4.4	Application to Insurance . . . . .	59
<b>5</b>	<b>Tools and Methodology — Estimating Onset Rates</b>	<b>60</b>
5.1	Onset Rate and Penetrance . . . . .	60
5.2	Classical Survival Analysis . . . . .	60
5.3	Pedigree Likelihoods . . . . .	62
5.3.1	Example . . . . .	63
5.4	Genetic Test Results . . . . .	65
5.5	Huntington’s Disease . . . . .	66
5.5.1	Mutation Carrier . . . . .	66
5.5.2	Non-Mutation Carriers . . . . .	67
5.6	Breast and Ovarian Cancer . . . . .	67
5.6.1	Mutation Carriers . . . . .	68
5.6.2	Non-Mutation Carriers . . . . .	70
5.7	Variance of Parameter Estimates . . . . .	71
5.8	Bootstrapping . . . . .	72
5.9	Ascertainment Bias . . . . .	73
<b>6</b>	<b>Tools and Methodology — Estimating Premium Rates</b>	<b>77</b>
6.1	The Markov Model . . . . .	77
6.2	The Model of Critical Illness Insurance Given Genetic Information .	79
6.3	Intensities . . . . .	79
6.3.1	Cancer . . . . .	80
6.3.2	Heart Attack . . . . .	81
6.3.3	Stroke . . . . .	81
6.3.4	28-day survival probabilities . . . . .	82

6.3.5	Other CI Illness Claims . . . . .	82
6.3.6	Mortality . . . . .	83
6.4	Calculating Extra Premiums when Genotype is Known . . . . .	83
6.5	The Insurance Model for Huntington’s Disease . . . . .	83
<b>7</b>	<b>Huntington’s Disease</b>	<b>87</b>
7.1	Data . . . . .	87
7.2	An Example of a Pedigree . . . . .	88
7.3	Missing Data Problems . . . . .	89
7.4	Results — Onset Rates . . . . .	92
7.5	Results – Premium Rates . . . . .	94
7.5.1	Ascertainment Bias . . . . .	105
7.6	Conclusions . . . . .	112
<b>8</b>	<b>Breast and Ovarian Cancer</b>	<b>114</b>
8.1	Data . . . . .	114
8.2	Results – Onset Rates . . . . .	115
8.3	Results – Premium Rates . . . . .	120
8.4	Ascertainment Bias . . . . .	130
8.5	Conclusions . . . . .	144
<b>9</b>	<b>Familial Adenomatous Polyposis</b>	<b>145</b>
9.1	Data . . . . .	145
9.2	How Might Screening Affect Insurance Risk . . . . .	145
9.3	Year of Onset of FAP . . . . .	146
9.4	Year of Birth . . . . .	147



9.5	Method of Diagnosis . . . . .	148
9.6	Screening and Cancer Risk . . . . .	149
9.7	Conclusions . . . . .	151
<b>10</b>	<b>Conclusions and Recommendations for Future Work</b>	<b>153</b>
10.1	Huntington's Disease . . . . .	154
10.1.1	Onset Rates . . . . .	154
10.1.2	Premium Rates . . . . .	154
10.2	BRCA1 related breast and ovarian cancer . . . . .	155
10.2.1	Onset Rates . . . . .	155
10.2.2	Premium Rates . . . . .	156
10.3	Effect of adjusting for ascertainment bias . . . . .	156
10.4	Implications for GAIC criteria . . . . .	157
10.5	Familial Adenomatous Polyposis . . . . .	158
10.6	Further Work . . . . .	159
10.6.1	Breast and Ovarian Cancer . . . . .	159
10.6.2	Huntington's Disease . . . . .	160
10.6.3	Familial Adenomatous Polyposis . . . . .	161
<b>A</b>	<b>Adult Polycystic Kidney Disease</b>	<b>162</b>
A.1	Epidemiological Review . . . . .	162
A.1.1	Genetic Testing . . . . .	163
A.1.2	Ultrasonographic Scan (USS) . . . . .	164
A.1.3	Age at Onset of ESRD . . . . .	165
A.1.4	Survival . . . . .	166
A.1.5	Gender Difference . . . . .	168

A.1.6 Hypertension . . . . .	168
A.2 Critical Illness Insurance Studies . . . . .	168
A.3 Available Data and Problems . . . . .	170
<b>B Maximum Likelihood Estimation</b>	<b>172</b>
B.1 Maximising the Likelihood . . . . .	172
B.1.1 Example . . . . .	173
<b>C Other Tools Used</b>	<b>176</b>
C.1 Numerical Integration . . . . .	176
C.2 Polynomial Approximation . . . . .	177
C.2.1 Normal Distribution . . . . .	177
C.2.2 Gamma Function . . . . .	177
<b>References</b>	<b>179</b>
<b>Bibliography</b>	<b>209</b>

# List of Tables

2.1	All possible codons and their corresponding amino acids (one-letter code). . . . .	7
2.2	The 1-letter codes for each amino acid as well as the name of each amino acid and the corresponding 3-letter code used. . . . .	8
4.1	Stages of progression of Huntington’s Disease for a typical patient. Source: Harper (1996). (Note: Chorea refers to brief, repetitive, jerky, or dancelike uncontrolled movements caused by muscle contractions.) . . . . .	39
4.2	Mean age at onset of HD in published studies. $N$ is the number of subjects in the study, $\mu$ is the mean age estimated by that study and $\sigma$ is the standard deviation. Source: Gutiérrez & Macdonald (2002a)	42
4.3	Correlation between CAG repeat length and age at onset of Huntington’s Disease in published studies. The Kehoe study published the results separately for those that inherited the mutation from their mother (maternal transmission) and father (paternal transmission).	43
4.4	Breast cancer penetrance in published studies. The penetrance is defined at age 70 for all studies except for Claus <i>et al.</i> (1998), Risch <i>et al.</i> (2001) and Anglican Breast Cancer Study Group (2000) where it is defined at age 80. . . . .	49

4.5	Ovarian cancer penetrance in published studies. The penetrance is defined at age 70 for all studies except for Claus <i>et al.</i> (1998), Risch <i>et al.</i> (2001) and Anglican Breast Cancer Study Group (2000) where it is defined at age 80. . . . .	50
4.6	Proportion of people with colorectal cancer at time of diagnosis, in the call-up group and in the propoiti group. . . . .	56
5.1	Non-zero terms in the likelihood (5.9) for the pedigree in Figure 5.2. Genotypes are $u_i = 1$ = unaffected and $u_i = 2$ = affected. $f$ is the frequency of the mutation within the population . . . . .	63
6.1	Proportions of the whole population who are in each of the three sub-populations. Source: Gui <i>et al.</i> (2006). . . . .	80
6.2	28-day mortality rates after a heart attack for males. Source: Dinani <i>et al.</i> (2000). . . . .	83
6.3	Proportions of the whole population who are in each of the three sub-populations. Source: Harper, Lim & Craufurd (2000) . . . . .	86
7.1	The assumptions made about missing or indeterminate data and how each assumption should affect the rate of onset. A1–A3 and B1–B3 are the labels referred to in Section 7.3. . . . .	92
7.2	The Normal model and sensitivity analyses. The parameters $\hat{m}$ and $\hat{\sigma}$ are the mean and standard deviation of the age at onset, respectively, and $N$ is the sample size. The sampling variance of each estimate is shown in parentheses. Assumptions A1–B3 are described in Section 7.3. . . . .	93
7.3	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for the OR class, for claims at Stage 2 of Harper’s progression. Using parameters obtained given assumptions A1 & B1, and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. . . . .	101

7.4	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for OR class, for claims at Stage 3 of Harper's progression. Using parameters obtained, given assumptions A1 & B1, and assuming 'earliest possible age at onset'. The table also includes the standard deviation of the estimated premium. . . . .	102
7.5	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 2 of Harper's progression. Using parameters obtained given, assumptions A1 & B1, and assuming 'latest possible age at onset'. The table also includes the standard deviation of the estimated premium. . . . .	103
7.6	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 3 of Harper's progression. Using parameters obtained given, assumptions A1 & B1, and assuming 'latest possible age at onset'. The table also includes the standard deviation of the estimated premium. . . . .	104
7.7	The minimum and maximum bootstrapped values of $m$ when adjusting for ascertainment bias, where $m$ is the mean of the age at onset. . . . .	105
7.8	The minimum and maximum of the bootstrapped values of $m$ carrying out the random sampling from the multivariate normal distributions, where $m$ is the mean of the age at onset. . . . .	105
7.9	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 2 of Harper's progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming 'earliest possible age at onset'. . . . .	108

7.10	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 3 of Harper's progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming 'earliest possible age at onset'.	109
7.11	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 2 of Harper's progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming 'latest possible age at onset'.	110
7.12	Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 3 of Harper's progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming 'latest possible age at onset'.	111
8.1	The penetrance estimates that we use in the estimation of the onset rates of breast and ovarian cancer and the studies they were obtained from.	115
8.2	The Gamma model and sensitivity analyses using the penetrance estimates of Lalloo <i>et al.</i> (2006). $\alpha$ and $\beta$ are the parameters of the Gamma distribution and $N$ is the sample size. The variance of each estimate is shown in parentheses. Assumptions A1–A3 are described in Section 7.3.	117
8.3	The Gamma model and sensitivity analyses using the penetrance estimates of Antoniou <i>et al.</i> (2006). $\alpha$ and $\beta$ are the parameters of the Gamma distribution and $N$ is the sample size. The variance of each estimate is shown in parentheses. Assumptions A1–A3 are described in Section 7.3.	118

8.4	The Gamma model and sensitivity analyses using the penetrance estimates of Antoniou <i>et al.</i> (2003). $\alpha$ and $\beta$ are the parameters of the Gamma distribution and $N$ is the sample size. The variance of each estimate is shown in parentheses. Assumptions A1–A3 are described in Section 7.3. . . . .	119
8.5	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimates. . . . .	126
8.6	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. . . . .	127
8.7	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Laloo <i>et al.</i> (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. . . . .	128
8.8	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Laloo <i>et al.</i> (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. . . . .	129

8.9	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband. . . . .	136
8.10	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband. . . . .	137
8.11	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands. . . . .	138
8.12	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands. . . . .	139



8.13	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband. . . . .	140
8.14	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband. . . . .	141
8.15	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands. . . . .	142
8.16	Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands. . . . .	143
A.1	Mean age at onset of ESRD for people with APKD. . . . .	166

# List of Figures

2.1	Three generations of a sample pedigree. Squares are males, circles are females, and a slash denotes death. Affected individuals are shown as filled squares/circles. The year of birth, death and other relevant information for each person is shown. Individuals are labelled X:Y, where ‘X’ (Roman) labels the generation and ‘Y’ (Arabic) the individual. $x_i$ represents the age of person $i$ at onset or censoring. .	16
3.1	A multiple state model for BRCA1/BRCA2-related breast and ovarian cancer in Critical Illness insurance. . . . .	27
3.2	A multiple state model for HD in Critical Illness insurance. . . . .	33
5.1	The two-state model for critical illness insurance. . . . .	62
5.2	Three generations of a sample pedigree. Squares are males, circles are females, and a slash denotes death. Affected individuals are shown as filled squares/circles. The year of birth, death and other relevant information for each person is shown. Individuals are labelled X:Y, where ‘X’ (Roman) labels the generation and ‘Y’ (Arabic) the individual. $x_i$ represents the age of person $i$ at onset or censoring. .	64
5.3	A multiple state model for HD in Critical Illness insurance. . . . .	66
5.4	A multiple state model for BRCA1/BRCA2-related breast and ovarian cancer in Critical Illness insurance. . . . .	68
6.1	A multiple state model for BRCA1/BRCA2-related breast and ovarian cancer in Critical Illness insurance. . . . .	78

6.2	A multiple state model for HD in Critical Illness insurance. . . . .	85
7.1	Three generations of a sample HD pedigree. Squares are males, circles are females, and a slash denotes death. Affected individuals are shown as filled squares/circles. The year of birth, death and other relevant information for each person is shown. Individuals are labelled X:Y, where 'X' (Roman) labels the generation and 'Y' (Arabic) the individual and $x_i$ represents the age at onset or censoring. . . . .	89
7.2	Distribution of age at onset of Huntington's Disease estimated from the pedigree data, assuming both 'earliest' and 'latest age at onset'. The figure also shows the distributions of age at onset of HD obtained assuming a Normal distribution for the mean and standard deviation found by Bell (1934) and Wendt <i>et al.</i> (1959). . . . .	94
7.3	Sampling distribution of the premium rate for a range of ages, assuming 'earliest possible age at onset' and an insurance term of 10 years. The scales of the $x$ and $y$ axes are not the same for all four plots. Top Left: Female with CI event defined as Harper's Stage 3. Top Right: Male with CI event defined as Harper's Stage 3. Bottom Left: Female with CI event defined as Harper's Stage 2. Bottom Right: Male with CI event defined as Harper's Stage 2. . .	97
7.4	Sampling distribution of the premium rate for a range of ages, assuming 'latest possible age at onset' and an insurance term of 10 years. The scales of the $x$ and $y$ axes are not the same for all four plots. Top Left: Female with CI event defined as Harper's Stage 3. Top Right: Male with CI event defined as Harper's Stage 3. Bottom Left: Female with CI event defined as Harper's Stage 2. Bottom Right: Male with CI event defined as Harper's Stage 2. . .	98

7.5	Sampling distribution of the premium rate for a range of policy terms, assuming ‘earliest possible age at onset’ and the age at which the policy is taken out is 20 years old. The scales of the $x$ and $y$ axes are not the same for all four plots. Top Left: Female with CI event defined as Harper’s Stage 3. Top Right: Male with CI event defined as Harper’s Stage 3. Bottom Left: Female with CI event defined as Harper’s Stage 2. Bottom Right: Male with CI event defined as Harper’s Stage 2. . . . .	99
7.6	Sampling distribution of the premium rate for a range of policy terms, assuming ‘latest possible age at onset’ and the age at the which policy is taken out is 20 years old. The scales of the $x$ and $y$ axes are not the same for all four plots. Top Left: Female with CI event defined as Harper’s Stage 3. Top Right: Male with CI event defined as Harper’s Stage 3. Bottom Left: Female with CI event defined as Harper’s Stage 2. Bottom Right: Male with CI event defined as Harper’s Stage 2. . . . .	100
7.7	The sampling distributions for the premium rate assuming either Harper’s Stage 2 or Stage 3 as the CI event, with and without adjustment for ascertainment bias. The scales are not the same for both plots. Left: The distributions for Females of age 20 for a policy term of 40 years. Right: The distributions for Females of age 30 for a policy term of 30 years. . . . .	107
8.1	Distributions of age at onset for Breast Cancer obtained using the penetrance estimates of Antoniou <i>et al.</i> (2006), Antoniou <i>et al.</i> (2003) and Lalloo <i>et al.</i> (2006); Left: Assuming ‘Earliest age at onset’ and Right: Assuming ‘Latest age at onset’. . . . .	120
8.2	Distributions of age at onset for Ovarian Cancer obtained using the penetrance estimates of Antoniou <i>et al.</i> (2006), Antoniou <i>et al.</i> (2003) and Lalloo <i>et al.</i> (2006); Left: Assuming ‘Earliest age at onset’ and Right: Assuming ‘Latest age at onset’. . . . .	121

8.3	Distribution of age at onset for Breast Cancer (left) and Ovarian Cancer (right). Both ‘earliest age at onset’ and ‘latest age at onset’ are shown, assuming the penetrance estimates obtained by Lalloo <i>et al.</i> (2006). The distributions used by Macdonald <i>et al.</i> (2003) were obtained by fitting truncated Gamma distributions to the estimated incidence rates of Ford <i>et al.</i> (1998) for BRCA1 carriers. . . . .	122
8.4	Sampling distribution of the premium rate for a range of ages, for an insurance policy term of 10 years. Top Left: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Top Right: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest age at onset’. Bottom Left: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Bottom Right: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘latest age at onset’. . . . .	124
8.5	Sampling distribution of the premium rate for a range of policy terms, given the age at which insurance was taken out was 20 years old. Top Left: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Top Right: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest age at onset’. Bottom Left: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Bottom Right: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘latest age at onset’. . . . .	125

8.6	Sampling distributions of the premium rate, for a female aged 50 seeking an insurance policy term of 10 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Top Right: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest age at onset’. Bottom Left: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Bottom Right: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘latest age at onset’.	132
8.7	Sampling distribution of premium rate, for a female aged 40 seeking an insurance policy term of 20 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Top Right: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest age at onset’. Bottom Left: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Bottom Right: Using the penetrance estimates of Lalloo <i>et al.</i> (2006) and assuming ‘latest age at onset’.	133

8.8	Sampling distribution of premium rate, for a female aged 30 seeking an insurance policy term of 30 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Top Right: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest age at onset’. Bottom Left: Using the penetrance estimates of Laloo <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Bottom Right: Using the penetrance estimates of Laloo <i>et al.</i> (2006) and assuming ‘latest age at onset’.	134
8.9	Sampling distribution of premium rate, for a female aged 20 seeking an insurance policy term of 40 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Top Right: Using the penetrance estimates of Antoniou <i>et al.</i> (2006) and assuming ‘latest age at onset’. Bottom Left: Using the penetrance estimates of Laloo <i>et al.</i> (2006) and assuming ‘earliest age at onset’. Bottom Right: Using the penetrance estimates of Laloo <i>et al.</i> (2006) and assuming ‘latest age at onset’.	135
9.1	Left: Probability of survival given ‘onset occurred before 1990’ and probability of survival given ‘onset in or after 1990’. Right: Probability of survival given ‘onset occurred before 2000’ and probability of survival given ‘onset in or after 2000’.	147

9.2	Left: Probability of survival given ‘born before 1950’ against the probability of survival given ‘born in or after 1950’. Right: probability of survival given ‘born before 1970’ against the probability of survival given ‘born in or after 1970’. . . . .	148
9.3	Left: Probability of survival given ‘born before 1950 and onset before age 50’ against the probability of survival given ‘born in or after 1950 and onset before age 50’. Right: Probability of survival given ‘born before 1970 and onset before age 30’ against the probability of survival given ‘born in or after 1970 and onset before age 30’. . . . .	149
9.4	Time to onset for those diagnosed through screening compared with those diagnosed through other methods. . . . .	150
9.5	States of those who have been confirmed to have FAP, through testing, diagnosis (screening) and colorectal cancer. $N$ is the number of people that have been in each state. The average age is given as the average age, on entering the state, of all people who have been in this state. . . . .	152
A.1	A multiple state model for APKD in Critical Illness insurance. . . .	169
B.1	Example of the Simplex Method for a one parameter function, in two dimensions. (a) is the original simplex. (b) is the ‘reflection’ step, (c) is the ‘expansion’ step and (d) and (e) are ‘contraction’ steps. .	175



# Chapter 1

## Introduction

In the past 20 years the study of human genetics has advanced considerably and the genes responsible for many disorders have been identified. As a result presymptomatic genetic tests have been developed that can determine whether a person carries specific disease-related mutations and, subsequently, give a measure of the risk of developing the disease.

Some rare genetic disorders, such as Huntington's Disease, are purely genetic, and the gene mutation that has been identified occurs in all people who have the disorder. There is a 50% chance that a person has the associated mutation if one of their parents is a carrier, so a genetic test will inform an individual at 50% risk if their actual risk of having the mutation is 100% or 0% and, therefore, if they have a very significant risk of developing the disorder or no risk whatsoever.

For some other diseases, such as breast cancer, the genes that have so far been identified do not explain all cases of the disorder. Therefore, although there is a higher risk of breast cancer if a particular mutation is present, there is still a high chance of developing breast cancer in the absence of a known mutation.

While genetic tests inform people of their risks of developing a disorder there are worries that a result which reveals a high risk will disadvantage them, particularly when buying life and critical illness insurance.

From this arises the question of whether insurance companies should have access to the results of genetic tests when determining premium rates.

Insurance companies are concerned that if they do not have access to genetic test results they could experience adverse selection, where the applicant can take advantage of information about risks to their future health that the insurer is not aware of. Would a person with an adverse genetic test result buy more insurance than they would have done before they knew the result?

Insurers already take account of family histories of certain disorders when determining premium rates. We want to find out what effect the extra information gained through (possibly) being able to use genetic test results would have on premium rates.

Premium rates are usually determined based on the insurers' own experience of certain factors, such as age and smoking habits. As yet, however, they do not have information available from their own experience in respect of genetic testing.

Since 1999 the insurance industry in the UK has observed a self-imposed moratorium on the use of genetic testing, enforced by the Association of British Insurers (ABI). The moratorium states that genetic test results cannot be used by an insurer unless the amount sought is over a certain threshold and even then the insurer can only use existing test results if they have been approved for use by the Genetics and Insurance Committee (GAIC).

Of the genetic disorders that are of interest to insurers we will investigate three, Huntington's disease (HD), BRCA1/2-related breast and ovarian cancer and Familial adenomatous polyposis (FAP). For both HD and BRCA1/2-related breast and ovarian cancer we will investigate critical illness insurance for known mutation carriers based on onset rates derived from real pedigree data. We also look at the effect that genetic testing has on FAP and the possible implications for critical illness insurance.

Adult Polycystic Kidney Disease (APKD) was originally on the list of genetic disorders regarded by the ABI to be important to the insurance industry but was removed as it is thought to no longer be of interest, in terms of genetic testing, due to the use of ultrasound in its diagnosis. We look at APKD, but not in the detail of the other three disorders.

Chapter 2 contains a basic introduction to genetics including a definition of DNA

and a description of the relationship between genes, DNA and chromosomes. The chapter then introduces possible methods of inheritance, including an outline of Mendel's laws of inheritance followed by a discussion of mutations, how they can occur, and the problems they may cause. There is also a brief introduction to pedigrees, the information that can be obtained to aid in genetics investigations, and the problem of ascertainment bias.

We discuss insurance and how genetics impacts on insurance in Chapter 3. This chapter starts by introducing the ideas behind insurance risk and illustrating how insurance premiums are determined. We then discuss the moratorium, imposed by the ABI, on the use of genetic testing in insurance and highlight GAIC's role in decisions made in relation to genetic testing. This chapter also includes a review of previous studies concerning how the genetic disorders considered here affect life and critical illness insurance.

Although work has previously been carried out to investigate the impact of three of the above genetic disorders (APKD, HD and BRCA-related breast and ovarian cancer) on insurance, these studies have had to rely on risk estimates published in the epidemiological literature for their premium calculations.

Typically the distribution of the age at onset is used when determining insurance premiums. The variances of the parameters of the distribution of age at onset, however, are not usually published and the data are never published and, as a result, a suitable measure of the reliability of the insurance premium cannot be obtained.

Here we have data in pedigree form, obtained by Jacki Needs and Professor Lyndsay Prior from the records of the Institute of Medical Genetics at the University of Wales, Cardiff.

We will use this data to make parameter estimates of the distribution of age at onset for each disorder and the variance matrix associated with these parameter estimates. The estimates of variability will mean that, using Monte-Carlo methods, we can estimate the reliability of the premium rates that would be charged to mutation carriers.

In Chapter 4 we review epidemiological studies concerning the three genetic disor-

ders of interest to us. This includes previous estimates of all quantities of interest to insurers as well as details of the genetics for each disorder.

Chapter 5 reviews methods used to estimate onset rates of genetic disorders based on families with a history of the associated disease. We describe the statistical methods used and related issues, namely likelihoods, estimation of the sampling variances, bootstrapping and ascertainment bias. Chapter 6 briefly discusses the actuarial methods used to compute premium rates for persons with a known mutation.

The results obtained, both rates of onset of disease, and premium rates, are discussed for each disorder separately. Chapter 7 concentrates on Huntington's Disease, while Chapter 8 on BRCA1-related breast and ovarian cancer.

We treat familial adenomatous polyposis (FAP) in a different way for two reasons; (a) the disorder is no longer considered important by insurers; and (b) the information available indicated that the likelihood methods used for the other disorders were not suitable. Chapter 9 discusses the data available for FAP and the limited aims of the analysis carried out. A brief discussion of APKD, introducing the epidemiology of APKD and previous insurance studies carried out, is given in Appendix A.

Chapter 10 discusses the conclusions we reached and their possible applications to insurance. We also discuss the possibilities for future work that can be carried out using the data available here. The bibliography in this thesis gives a list of papers, books and websites which I found useful in researching the subjects within this thesis but have not referenced anywhere.

# Chapter 2

## Genetics

### 2.1 An Introduction to Genetics

According to the Human Genome Project (HGP) the number of genes in the human genome is much lower than previously thought. They estimate that every human has approximately 20,000 – 25,000 genes (International Human Genome Sequencing Consortium, 2004). These genes control the make-up of every part of the human body from blood type to hair colour. In order to understand what genes do it is important first to look at what genes are.

Genes are the units of inheritance for every living thing, carrying information from one generation to the next. They are made up of strands of genetic material called deoxyribonucleic acid (DNA) and are contained in the nucleus of every cell.

For more information on the basics of human genetics than we can give, we refer the reader to Strachan and Read (1999) and Pasternak (1999).

#### 2.1.1 DNA

Watson and Crick (1953) discovered that the DNA molecule is a double helix, shaped like a twisted ladder, and that it is made up of three main parts:

- Five-carbon sugars (deoxyribose).
- Phosphate molecules.
- Four nucleotide bases;
  - Adenine (A)
  - Guanine (G)
  - Cytosine (C)
  - Thymine (T).

The ‘sides’ of the ladder are made up of the sugars and phosphates, while the ‘rungs’ are made up of pairs of the bases linked to each other. A is always paired with T, and G is always paired with C. These four letters, A, G, C and T are the alphabet with which the message in DNA is encoded.

Through a process called transcription the DNA ‘ladder’ is split and a single strand, one half of the ladder, of DNA is transcribed into a single-stranded Ribonucleic Acid (RNA) polymer called mRNA (messenger RNA).

RNA is an important step between DNA and the manufacture of proteins. RNA, like DNA, is made up of a chain of nucleotides and also consists of a four-letter alphabet, but with thymine (T) replaced by uracil (U).

The mRNA may be thought of as a chain of three letter ‘words’ called codons. Each codon provides the code for an amino acid which is added sequentially to a protein chain. The resulting proteins control nearly all aspects of the human body.

There are 64 possible codons but only twenty possible amino acids so some amino acids are coded for by more than one codon. The start codon (AUG), signals the beginning of the coding sequence, while three stop codons signal the end (UAA, UAG and UGA). Table 2.1 shows the possible codons and the corresponding one-letter amino acid code. Table 2.2 shows what amino acid each one-letter code represents.

Essentially, through transcription, a segment of the DNA ladder is translated into a chain of 3-letter RNA words which, in turn, are translated into a chain of amino acids to form a protein.

Table 2.1: All possible codons and their corresponding amino acids (one-letter code).

	<b>U</b>	<b>C</b>	<b>A</b>	<b>G</b>
<b>U</b>	UUU (F) UUC (F) UUA (L) UUG (L)	CUU (L) CUC (L) CUA (L) CUG (L)	AUU (I) AUC (I) AUA (I) AUG (M)	GUU (V) GUC (V) GUA (V) GUG (V)
<b>C</b>	UCU (S) UCC (S) UCA (S) UCG (S)	CCU (P) CCC (P) CCA (P) CCG (P)	ACU (T) ACC (T) ACA (T) ACG (T)	GCU (A) GCC (A) GCA (A) GCG (A)
<b>A</b>	UAU (Y) UAC (Y) UAA (X) UAG (X)	CAU (H) CAC (H) CAA (Q) CAG (Q)	AAU (N) AAC (N) AAA (K) AAG (K)	GAU (D) GAC (D) GAA (E) GAG (E)
<b>G</b>	UGU (C) UGC (C) UGA (X) UGG (W)	CGU (R) CGC (R) CGA (R) CGG (R)	AGU (S) AGC (S) AGA (R) AGG (R)	GGU (G) GGC (G) GGA (G) GGG (G)

Table 2.2: The 1-letter codes for each amino acid as well as the name of each amino acid and the corresponding 3-letter code used.

1-Letter Code	Amino Acid	3-Letter Code
A	Alanine	Ala
C	Cysteine	Cys
D	Aspartic acid	Asp
E	Glutamic acid	Glu
F	Phenylalanine	Phe
G	Glycine	Gly
H	Histidine	His
I	Isoleucine	Ile
K	Lysine	Lys
L	Leucine	Leu
M	Methionine	Met
N	Asparagine	Asn
P	Proline	Pro
Q	Glutamine	Gln
R	Arginine	Arg
S	Serine	Ser
T	Threonine	Thr
V	Valine	Val
W	Tryptophan	Trp
Y	Tyrosine	Tyr
X	Stop Codon	



## 2.1.2 Chromosomes

DNA is organised into structures called chromosomes, each containing hundreds of genes. Humans have 23 pairs of chromosomes, 22 of which (numbered 1 – 22) are called autosomes while the 23rd pair are the sex chromosomes; one X and one Y chromosome in males and two X chromosomes in females. One chromosome of each of the 23 pairs is inherited from each parent.

Each pair of autosomes is similar in that they contain the same genetic information in the same order. They can, however, differ in the specific form of the genes. These different forms of a specific gene, given by different nucleotide sequences within the gene, are called alleles.

The observed allele at a genetic locus is called the genotype, while the outward appearance, characteristic or trait, influenced by the genotype is known as the phenotype. Different genotypes can result in different phenotypes in individuals, for example the different eye colour genotypes can give the phenotypes blue, green and brown amongst others.

A person who has two different alleles at a genetic locus is known as heterozygous, while a person with identical alleles at a genetic locus is known as homozygous.

The wild-type allele is the common form of the allele in the population, often considered to be that which gives the ‘normal’ phenotype.

## 2.2 Inheritance

### 2.2.1 Cell Division and Reproduction

To understand how genes are inherited we need to look at the mechanisms of cell division and reproduction, namely mitosis and meiosis.

**Meiosis** Everyone inherits one copy of each chromosome from their mother and the other from their father. The make-up of the actual chromosomes inherited from the parents is determined during meiosis.

Meiosis is the process through which the chromosomes to be passed on are installed in gametes (the reproductive egg cells, in the mother, and the reproductive sperm cells, in the father). Each gamete contains 23 chromosomes, instead of 23 pairs of chromosomes.

During meiosis a mixture of the genes from the maternally and paternally inherited chromosomes is produced. The resulting chromosomes in the gametes contain genetic information from both the mother and the father of the parent of the prospective child.

**Mitosis** Mitosis is the process of cell division. The division of the ‘mother’ cell, through mitosis, leads to two ‘daughter’ cells, where each daughter cell has a set of chromosomes identical to that of the mother cell.

This is the process through which the body grows, from the fertilised egg (zygote), to a fully grown adult.

## 2.2.2 Mendel’s Laws of Inheritance

Gregor Mendel derived his laws of inheritance while conducting experiments on plant hybrids (Mendel, 1866, in Peters, 1973). These laws relate to the transmission of hereditary characteristics and traits from parents to their offspring.

The laws are as follows;

### (1) Mendel’s Law of Dominance

When two people have a different trait, their offspring will only display the dominant trait in the phenotype. For example, hair colour; if a child inherits the blond genotype from one parent and the brown from the other, the child will have brown hair as this trait is dominant.

### (2) Mendel’s Law of Segregation

Each parent has two copies of each gene for each characteristic, inherited from their parents, and one copy of each gene is passed on to their offspring, through gametes.

### (3) Mendel's Law of Independent Selection

The emergence of one trait will not affect the emergence of another.

As well as applying to the inheritance of characteristics such as eye colour, hair colour and blood type, these laws also apply to the inheritance of mutations and genetic disorders.

## 2.2.3 Mutations

A mutation is a change in the DNA which can cause changes in the make-up of a person. These can be beneficial, neutral or harmful and can take many forms.

Here we will concentrate only on those mutations that are of interest to the insurance industry, that is, those that result in a high risk of developing disease among previously asymptomatic healthy adults.

Mutations can be classified in a number of ways:

### (a) By effect on the structure of the protein

- Small scale mutations.

- \* Point Mutations – These cause a single nucleotide to be exchanged for another.

- Missense mutations – These result in a change in the amino acid coded for by that codon. The change to the protein may be enough to cause it to be structurally abnormal or unstable.

- Nonsense mutation – This mutation alters an amino acid codon to a stop codon resulting in a truncated protein. The change either results in a protein with altered function or a very unstable protein.

- \* Deletion or Insertion – The deletion or insertion of a number of nucleotide bases.

- Frameshift mutation – This occurs when the number of deleted or inserted base pairs is not a multiple of three. After the point

of mutation all of the amino acids that follow will be different. Either an abnormal protein or no protein is produced.

- In-Frame mutation – This occurs when the number of deleted or inserted base pairs is a multiple of three. An unexpected amino acid will appear in, or disappear from, the protein. The protein may still be able to function.
- Large scale mutations.
  - \* These mutations are normally associated with congenital birth defects or can result in the foetus dying before birth.

#### **(b) By effect on function**

- Loss of function mutations – The resulting gene product has less or no function. These mutations normally have recessive phenotypes.
- Lethal mutations – The resulting gene product gives a phenotype which results in the development of a lethal disorder. This includes congenital, early-onset and late-onset disorders.

#### **(c) By aspect of phenotype affected**

- Morphological mutations – Changes in the appearance of individuals, such as, height, hair colour etc.
- Biochemical mutations – Disruptions or loss of biochemical functions. A mutation in the pathway which converts amino acid tyrosine to skin pigment melanin results in albinism.

Mutations can occur during meiosis or mitosis or can be caused by molecular decay and can be induced through chemical action and radiation.

### **2.2.4 Disorders**

Mutations can result in the development of disorders, from the relatively harmless colour-blindness to late-onset lethal disorders such as Huntington's Disease. These disorders can be classified in the following way.

(a) Monogenic (single-gene) disorders are those that are caused by a mutation in a single gene, or pair of genes at a single locus. There are a number of ways in which a monogenic disorder can be inherited, but here we will describe only two, to highlight the differences and similarities:

- *Autosomal Dominant Inheritance*

- Only one copy of a disease allele is needed for a person to develop the disease.
- There is a 50% chance of any offspring inheriting the mutation from an affected parent.
- The proportion of affected males and females should be equal.
- An example of an Autosomal Dominant disorder is Huntington's Disease.

- *Autosomal Recessive Inheritance*

- Two copies of a disease allele are needed for a person to develop the disease.
- The parents of an affected person are usually unaffected carriers.
- There is a 25% chance of a child, whose parents are unaffected carriers, inheriting two copies of the mutation and having the phenotype.
- There is a 25% chance of a child, whose parents are unaffected carriers, inheriting no copies of the mutation.
- There is a 50% chance of a child, whose parents are unaffected carriers, inheriting one copy of the mutation and being a carrier.
- The proportion of affected males and females should be equal.
- An example of an Autosomal Recessive disorder is Cystic Fibrosis.

Genetic heterogeneity may exist in some monogenic disorders; this occurs when mutations at more than one locus or different mutations at the same locus can result in the development of the disorder.

(b) Polygenic (multifactorial or multi-gene) disorders are those that are caused by variations in a combination of genes.

- The inheritance of the mutations in several genes may make a person more susceptible to a disease.
- Some environmental factors may combine with genotype to contribute to the disease.
- An example of a possible polygenic disorder is Coronary Heart Disease.

It is to be expected that identifying disease-causing mutations will be easier in respect of monogenic disorders than of polygenic disorders.

A number of genetic disorders do not develop until later ages, late-onset disorders. These late-onset disorders are usually associated with mutations that cause tissue deterioration over time, or with mutations in the DNA repair genes that can result in cancer developing. The diagnosis of these disorders does not usually occur until the deterioration has reached some limit and the symptoms become evident, as is the case with Huntington's Disease.

## 2.3 Pedigrees

In the context of hereditary disease the family tree is called a pedigree. The pedigree enables the tracking of disease through generations of the family. Figure 2.1 shows an example of a pedigree.

Most pedigrees are laid out in the same conventional way, with the youngest generation at the bottom and working back as far as possible to the pedigree founders. Males are symbolised by squares while females are symbolised by circles. A slash through either symbol indicates that that individual is dead and a filled symbol indicates onset of the disease with which the pedigree is concerned.

The ideal pedigree would have dates of birth, death, onset and any genetic tests, for every individual, but unfortunately this is often not the case in practice.

Pedigrees are compiled from information provided to geneticists or researchers. Each pedigree usually comes to the attention of geneticists through one or more affected people, often referred to as probands or index cases. In the case of HD it usually only needs one person to be diagnosed with the disorder to highlight the

pedigree as the disorder is rare and purely genetic.

The risk of developing breast or ovarian cancer, however, is associated with factors other than the two mutations that have been identified so far. As a result it is often the case that pedigrees which carry either mutation are only highlighted through multiple cases of BC and/or OC occurring within them.

### **2.3.1 Information That Can Be Obtained From Pedigrees**

Pedigrees are very useful to investigators of hereditary and genetic diseases. They enable them to look at the patterns of inheritance of diseases and traits.

A number of quantities are important to clinicians and researchers investigating genetic diseases, including; the onset rate, the post-onset survival rate and the mode of inheritance. These quantities can be estimated through the examination of pedigrees of families who display the disease.

Survival analysis usually aims to estimate probabilities of events such as; survival from birth, survival after development of a disease, or survival after treatment. The estimates can be used to compare the effect that different factors, such as different treatments and different types of the disease, have on the probability that someone lives to a certain age.

An onset rate is similar, but instead of survival to death, we can think of it as survival to onset of the disorder. Many studies concerning onset of a disease publish estimates of the median age at onset and sometimes Kaplan-Meier curves for survival to onset.

Often genetic and epidemiological studies publish an estimate of the penetrance of the disease, equivalent to the onset rate. The penetrance at age  $x$  is defined as the probability that onset has occurred by age  $x$ . Some genetic disorders are completely penetrant, which means that everyone with the mutation will develop the disease, if other decrements do not occur first. Other disorders may have incomplete penetrance, meaning that not all people with the mutation will develop the disease in their lifetime, even if all other decrements were absent.

A more detailed definition of onset rate and penetrance is given in Section 5.1, while

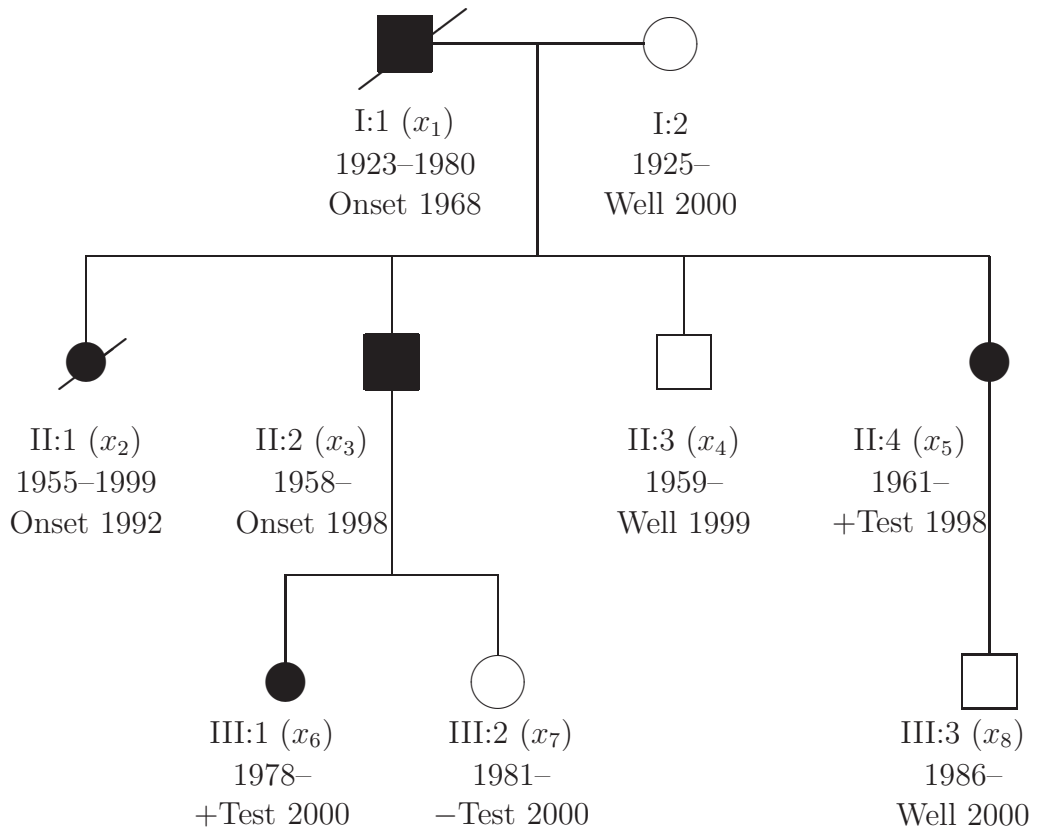


Figure 2.1: Three generations of a sample pedigree. Squares are males, circles are females, and a slash denotes death. Affected individuals are shown as filled squares/circles. The year of birth, death and other relevant information for each person is shown. Individuals are labelled  $X:Y$ , where ‘ $X$ ’ (Roman) labels the generation and ‘ $Y$ ’ (Arabic) the individual.  $x_i$  represents the age of person  $i$  at onset or censoring.



Sections 5.3 – 5.6 describe the methods used in this investigation to estimate them. Chapter 4 discusses previous work carried out concerning the genetic diseases of interest here and, where relevant, gives estimates of the onset rates obtained in these studies.

### **2.3.2 Ascertainment Bias**

When carrying out investigations of rare autosomal dominant disorders it is not sensible to attempt to obtain a study sample by sampling from the whole population and testing for the associated mutations. What is more likely to happen is that, as noted in Section 2.3, families who carry a mutation associated with a particular disease will be identified by geneticists and researchers through one or more affected people, namely ‘probands’.

Obtaining a sample in this way results in a bias, usually known as ascertainment bias, as the pedigrees used were not selected at random from the population. Ascertainment bias is a major issue when carrying out studies concerning genetic disorders and some attempt should be made to account for it.

Section 5.9 provides more detail on ascertainment bias and discusses some of the possible methods that can be used to adjust for it. Some attempt is usually made to adjust for ascertainment bias in epidemiological studies concerning genetic disorders. However, there is often a lack of information about the effect of ascertainment bias on the parameter estimates. Usually only the estimates, such as the penetrance or the parameters of the distribution of age at onset, obtained after adjustment has been made, are published.

It is often of interest in insurance studies to examine how ascertainment bias, and any subsequent adjustment for it, may affect the resulting insurance premiums. As the literature does not routinely publish estimates with and without adjustment for ascertainment bias and, as the pedigree data used to obtain the estimates of penetrance etc are not given, some insurance studies attempt to adjust for ascertainment bias by adjusting the penetrance estimates. Macdonald *et al.* (2003), for example, reduced the onset rates by 50 and 70% so as to determine the possible effects of ascertainment bias on premium rates by assuming that the effect of ascertainment

bias on onset rates was very large.

As this study makes use of original pedigree data we can make some attempt to adjust for ascertainment bias and, in turn, evaluate the effect the adjustments have on the resulting premium rates. Section 5.9 identifies the method we use to adjust for ascertainment bias.

# Chapter 3

## Insurance

### 3.1 Insurance and Risk

When taking out motor insurance a number of questions will be asked to determine how much risk a person and their car present. The main questions relate to age, number of years' driving experience and number of previous claims, as well as where the applicant lives and how much their car is worth.

All of these are understandable; younger, less experienced drivers are known to be more likely to have an accident, while home address and the location where the car is parked may affect the risk of a car being stolen, and the value of a car will affect how much might be paid out.

In respect of life and critical illness insurance the questions are mainly health related. The main risk factors are age (the risk of death increases as an individual ages) smoking status (smokers are at a higher risk of death than non- or ex-smokers) and sex (males are at higher risk of death and disease at younger ages though their risk at later ages does tend to be similar to that of females).

Also taken into account is the state of the applicant's own health as well as that of their family, called 'family history'. Family history is information about the state of the applicant's family's health including any cases of diseases such as heart disease, stroke and cancer.

The insurer considers all relevant factors, estimates a person's risk of death or of

suffering a critical illness and consequently determines the premium rate they should pay. Normally in insurance there are three main categories of risk.

The main class is called the ordinary rates (OR) class. This accounts for about 97% (Leigh, 1990) of applications for life insurance. The OR is the rate paid by people who are in good health, although it does include some people who have slightly poorer than average health. The OR class for life insurance is estimated to include people with up to 130% to 150% of the average rates of mortality (Leigh, 1990).

If the applicant is estimated to have a higher rate of morbidity or mortality than can be included in the OR class, because of poor health or adverse family history, they may be charged a higher premium, usually expressed as a multiple of the OR premium. This class accounts for about 2% of the life insurance population (Leigh, 1990). The final 1% of applicants are declined life insurance as the premium they would be charged exceeds about 500% of the OR premium (Macdonald, 2000) for life insurance. For critical illness insurance if the premium an applicant would be charged exceeds 300% – 350% of the OR premium they would be declined cover (Gutiérrez & Macdonald, 2003). Dinani *et al.* (2000) states that 15% of all critical illness insurance claims are declined, 22% of which are declined due to non-disclosure and 70% due to the definitions of critical illness not being met.

## 3.2 Use of Genetic Information by UK Insurers

Genetic information can be defined as:

- A family history of a Mendelian disorder such as Huntington’s Disease.
- A family history of a complex non-Mendelian disorder such as heart disease.
- The results of tests such as those relating to blood, urine and ultrasound which identify a genetic disease and as a result confirm the existence of a mutation.
- The result of a DNA-based test.

Since the development of DNA-based tests for some genetic disorders, concerns have been raised about who should have access to the results. One of the main questions

is whether insurance companies should be able to use presymptomatic genetic test results to calculate premiums. It is thought that some people may be reluctant to undergo presymptomatic testing because they are worried that an adverse result will mean that they are uninsurable, or that people with an adverse result may not attempt to buy insurance as they believe they would not be offered cover. These problems have to be taken very seriously by the insurance industry.

In 1997 the Association of British Insurers (ABI), set out a code of conduct for insurers to follow concerning the use of genetic information. They began by defining genetic tests as those that examine DNA or RNA only. They also imposed a five year moratorium, which has since been extended until 2011, on the use of presymptomatic genetic test results (ABI, 2005).

The most recent moratorium of the ABI stated that genetic test results should not be used except, possibly, in respect of policies over £500,000 for life insurance and £300,000 for CI insurance. Even if the applicant wishes to buy insurance above these levels, they cannot be asked to take a presymptomatic genetic test and the result of an existing genetic test can only be used if its use has been approved by the Genetics and Insurance Committee (GAIC).

GAIC was formed by the UK Government in 1999 to oversee the use of genetic testing in the insurance industry. It is GAIC's job to decide whether a genetic test result can be used by an insurer, following an insurer's application. In practice the ABI, rather than individual insurance companies, applies for permission to use specific genetic tests. Applications are made for specific genetic tests and specific insurance products and must contain the following information:

- Evidence that the genetic test is accurate and reliable.
- Clinical evidence that the the presence of a mutation leads to implications for the health of an individual.
- Actuarial evidence that there is an increased probability of the insured event occurring if the mutation is present.

GAIC's annual report (2000) defines the evidence required in more detail. The ABI developed a list of eight genetic disorders that their genetics adviser defined

as important to insurers as part of their code of conduct (ABI, 1997). The eight disorders are listed below along with a brief description of each disease and the mutations that are known to cause them.

**Huntington's Disease (HD).** HD is a progressive neurological disorder which affects movement and speech. It is a fully penetrant, purely genetic, disorder which is caused by a mutation in a single gene, the huntingtin gene.

**Familial Adenomatous Polyposis (FAP).** FAP is a disorder affecting the colon and rectum, through the development of polyps which, if left untreated, may become cancerous. A mutation in the APC gene is the main cause of FAP, although it is believed that other mutations result in FAP, including an autosomal recessive mutation, MUTYH.

**Breast Cancer & Ovarian Cancer (BC & OC).** Mutations in two genes have been found to cause breast and ovarian cancer, namely, BRCA1 and BRCA2. These mutations do not account for all cases of BC & OC and it is thought there are mutations still to be discovered. The presence of a mutation increases the risk of developing BC and OC significantly.

**Adult Polycystic Kidney Disease (APKD).** APKD is a disorder that results in cysts developing in the kidneys and leads to the need for dialysis and subsequently kidney transplantation. Two APKD genes have been identified, PKD1 and PKD2. These account for the majority of cases of APKD but it is believed that a third mutation may exist.

**Myotonic Dystrophy (MD).** MD is a muscle wasting disease caused by a mutation in the MDPK gene.

**Multiple Endocrine Neoplasia (MEN).** MEN results in the development of endocrine tumours. There are two types of MEN associated with different mutations. MEN1 is the only known gene associated with type 1, while RET is the only known gene to be associated with type 2. The ABI only considered the latter on their list.

**Hereditary Motor & Sensory Neuropathy (HMSN).** In this disorder the nerves to the muscles, particularly in the lower legs and hands, do not work

properly. There may be some effect on sensation in the affected areas as well. There are a number of genes associated with HMSN but the most prevalent is PMP22. The test for this gene was the only one considered by the ABI.

**Early-Onset Alzheimer's Disease (EOAD).** EOAD is a disease which results in progressive dementia. Three EOAD mutations have been identified, APP, PS1 and PS2, although these do not account for all cases.

This list was later amended to include only seven disorders (APKD was removed due to the use of ultrasound in its diagnosis).

To date only one application to GAIC, for Huntington's Disease and life insurance, has been successful. It was announced in October 2000 (Daykin *et al.*, 2003). The ABI intends to make other applications, concerning CI and income protection (IP) insurance for HD and CI, IP and life insurance for BRCA-related Breast and Ovarian Cancer.

The ABI has not yet considered applications to GAIC for the other disorders (EOAD, FAP, MEN, MD and HMSN). They decided that, based on factors such as age at onset and insurers' experience of purchasing patterns, the tests for these disorders may not be of practical value at the moment.

We give a more detailed discussion of the disorders which we consider (HD, FAP, and BC & OC) in Chapter 4 (and discuss APKD in Appendix A).

### **3.3 Information Available About CI and Life Insurance Risks**

When a person applies for life or CI insurance they are asked questions about a number of aspects of their life including; their health, their family history and their lifestyle. It is then the job of the underwriter to examine the mortality and morbidity risks associated with the information given and to determine whether the applicant can be insured at the ordinary rate.

The underwriter may need more information in order to reach a conclusion. They

may request a General Practitioner's Report (GPR) or a Medical Examination Report (MER) which will give them information about the applicant's present health, through physical measurements and non-DNA-based tests such as; height, weight, blood pressure, blood and urine tests and X-rays.

The underwriter then uses all available information to determine whether or not the risk presented by an applicant is acceptable, and to classify applicants with an acceptable risk to be in the OR class or in a substandard class. The risk classification is carried out so as to protect the insurer, by declining the people with severe risk and charging those at significantly increased risk higher premiums. It also aims to protect the other policyholders, who will ultimately lose out if the insurer admits a significantly mis-priced group of risks.

The risk classification is usually carried out using data already available, such as published epidemiological studies or life tables based on insurers' own experience. The data life insurers obtain through their own insurance experience is often more reliable because of the size of the sample and because the insured population usually has much lighter mortality than the general population.

There are currently no industry standard tables for CI insurance, as it is a relatively new insurance product, but the Continuous Mortality Investigation (CMI), who carry out research into mortality and morbidity experience, have begun an investigation and have published a number of working papers on the subject. They say that the CI insurance experience is, as yet, too immature for the data to allow the construction of standard CI tables (CMI Working Papers 14 & 18).

When an insurer does not ask questions about a possible risk factor it is impossible for it to base any assessment of the risk associated with that factor on their own experience. As a result of the moratorium on presymptomatic genetic tests, insurers have no information of their own about the risks associated with a positive result. Insurers also have little information about the risks associated with positive presymptomatic tests as genetic testing is relatively new and is not commonly used yet. There are, however, epidemiological studies that provide estimates of the rates of onset or death for some of the most important genetic diseases. Some of these studies will be discussed in Chapter 4.



## 3.4 Adverse Selection

Adverse selection occurs when a person attempts to buy insurance while knowingly concealing some relevant information. The applicant may try to get more insurance than they would normally, or even just try to get insurance when they would not otherwise have considered it.

For example an applicant may know they have heart disease but may deliberately not inform the insurer. This affects the insurer as they will evaluate the risk based on the wrong information, thus the claim costs will be higher than expected and the cost will be forced up for all policyholders.

A lot of the scope for adverse selection is eliminated through the insurer using medical and family history information. Also, the fact that deliberate non-disclosure may void an insurance policy should deter people from trying to defraud an insurer.

Now, however, insurance companies are worried that, with the increasing availability of genetic tests, adverse selection will increase. It is possible that a person could buy more insurance, given a positive result, or less insurance, given a negative result.

As the moratorium prevents insurers from asking about presymptomatic genetic test results they can only determine the presence of genetic risk from family history and, possibly, non-DNA-based medical tests. For example APKD manifests itself through cysts in the kidneys and can be detected using ultrasound.

We would like to know whether adverse selection, related to genetic testing, actually does occur, and if so how much effect it has. There has been some investigation of adverse selection related to genetic testing in general, as well as some studies about specific genetic disorders. Macdonald (1997) produced some crude estimates given some very extreme assumptions about genetic testing. He stated that:

“If insurance companies refrain from using (or are forbidden to use) the results of any genetic test in underwriting, additional mortality costs are likely to arise. However, if adverse selection does not extend to untypically large sums assured the magnitude of these costs is greatly reduced; large sums assured is the costliest aspect of adverse selection”

Macdonald (1999) concluded that if adverse selection involving higher sums assured can be avoided then a reasonable order of magnitude for the additional costs of adverse selection is 10%. He also demonstrated that it is unlikely that a separate risk class will emerge for those undergoing genetic tests.

### 3.5 Research Concerning Genetics and Insurance

Some work has been carried out, using published epidemiological studies, to give point estimates of premium rates for critical illness and/or life insurance. The following is an outline of such work in respect of critical illness insurance, for the genetic disorders that are of interest here.

Many of these papers discuss the effects on large and small insurance markets, where the large insurance market defines insurers that are well established and conversely the small insurance market defines insurers that are newer.

#### 3.5.1 Breast and Ovarian Cancer

Figure 3.1 shows a multiple state model for critical illness insurance for BRCA1 and BRCA2 related breast and ovarian cancer based on the APKD CI insurance model, introduced by Gutiérrez & Macdonald (2003) and discussed in Appendix A. The population is divided into sub-groups (usually three but sometimes more):

- Those with no family history, who are not at risk.
- Those at risk because of a family history but who do not carry a mutation.
- Those at risk because of a family history and who do carry a mutation.

The transition intensities in Figure 3.1 have to be estimated for each of these sub-groups. The rates of onset of breast cancer,  $\mu_{01}(x)$ , and ovarian cancer,  $\mu_{02}(x)$ , are usually based on estimates derived by epidemiological studies. The other intensities,  $\mu_{03}(x)$  and  $\mu_{04}(x)$  are estimated using a variety of medical and demographical studies (discussed in more detail in Section 6.3)

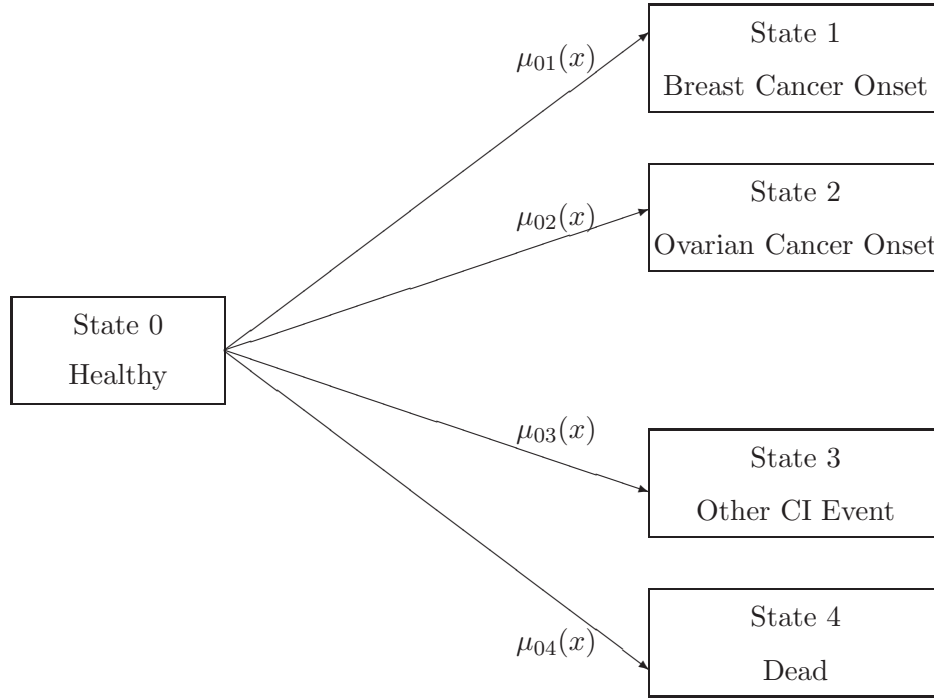


Figure 3.1: A multiple state model for BRCA1/BRCA2-related breast and ovarian cancer in Critical Illness insurance.

### Macdonald, Waters & Wekwete (2003a & 2003b)

Macdonald, Waters & Wekwete Part I (2003a) aimed to estimate the probability of an applicant having a BRCA1 or BRCA2 mutation given their family history, or more specifically, the structure of their family and the joint genotype of the family.

The population incidence rates of breast and ovarian cancer were estimated based on cancer registrations in England and Wales in 1990–92 (O.N.S., 1999), and mid-year population estimates in 1989–93. The breast and ovarian cancer onset rates ( $\mu_{01}(x)$  and  $\mu_{02}(x)$ ) among mutation carriers were determined from estimates in Ford *et al.* (1998).

These intensities were then used to determine the probabilities of having a BRCA1 or BRCA2 mutation given the probability of different family structures (and allowing for different mutation frequencies). The authors stated that risk estimates based on studies of high-risk families should perhaps not be applied to other populations.

Part II (2003b) investigated a model of critical illness insurance which incorporates breast and ovarian cancer as distinct causes of a claim, given the intensities and

probabilities estimated in Part I.

The authors used their model to estimate the costs of critical illness insurance for specific genotypes and found that level extra premiums were very high and that most applicants would not have been insurable.

The critical illness insurance premiums given a family history of breast and ovarian cancer were also estimated. The premiums were found to be significantly lower if there is a family history but no known genetic test result. They also found that the extra premiums varied depending on the age, term of policy, and what was known about the family history and the family structure.

An attempt was made to adjust for ascertainment bias by reducing the penetrance estimate. This was found to affect the costs of critical illness insurance significantly; the extra premiums fell significantly for both the genotype-specific and family history models. The costs of adverse selection were found to be high in small CI insurance markets. This estimate, however, was dependent on a number of extreme assumptions, therefore the authors suggested that the costs they estimated could be overstated.

### **Gui *et al.* (2006)**

This study introduced family history into the life history of an applicant. This was done by including the development of a family history as a transition between states in the model. Most breast and ovarian cancers are caused by factors other than the mutations that have been identified. Therefore the onset of either cancer alone does not identify a mutation carrier and does not give information in a simple way about the risks of family members being carriers. Also as being a mutation carrier does not always result in the onset of BC or OC the model of onset for BRCA-related BC and OC is more complex than that of fully penetrant, purely genetic disorders such as HD.

The post-onset mortality rates for breast and ovarian cancer were determined from Coleman *et al.* (1999). The rates of onset of breast and ovarian cancer for BRCA1 and BRCA2 mutations were based on the meta-analysis of Antoniou *et al.* (2003). Antoniou *et al.* (2003) used data from a number of studies to determine onset rates of breast and ovarian cancer. They only used studies where the index case was

identified independently of family history and had suffered breast or ovarian cancer. Adjustment for any possible ascertainment bias was made by maximising the conditional likelihood of the pedigree given the phenotypic and genotypic information of the index case.

The population was divided by Gui *et al.* (2006) into the following five sub-populations to incorporate family history;

- Family with no BRCA1 or BRCA2 mutation carriers.
- BRCA1 mutation carrier family but applicant not a carrier.
- BRCA1 mutation carrier family and applicant is a carrier.
- BRCA2 mutation carrier family but applicant is not a carrier.
- BRCA2 mutation carrier family and applicant is a carrier.

The underwriting threshold for a family history was defined as two first-degree relatives suffering onset of breast or ovarian cancer before the age of 50. Assumptions about the family structure were also made; all the applicant's siblings were supposed to be the same age as she, and their mother was supposed to be 30 years older. These assumptions meant that the family structure was defined by the number of siblings. The rate of onset of a family history was then estimated for each of the sub-populations.

The premium rates for life insurance for mutations carriers were, again, found to be significantly lower when adjustment was made for ascertainment bias (by again reducing the penetrance estimate). The life insurance premium rates for those with a family history were below the underwriting threshold, of 400 – 500% of the OR class.

It was found that the costs of adverse selection as a result of a moratorium on all genetic test results would have the greatest impact on the small insurance market. The percentage increases in premium rates, under severe adverse selection and a normal rate of insurance purchasing, were estimated to be 0.000219 for small insurance markets and 0.000510 for large insurance markets. A moratorium on ge-

netic test results and family history would result in premium rates increasing by 0.013602% and 0.001677%, respectively.

They applied the model to CI insurance and found that the premium increases were much higher than those for life insurance. Most mutation carriers would be uninsurable, while all people with a family history would be insurable.

The adverse selection costs for critical illness insurance were slightly higher than, but similar to, those for life insurance, under both moratoria.

### **Macdonald & McIvor (2006)**

Although BRCA1 and BRCA2 mutations carry a very high risk of developing breast and ovarian cancer they only account for about 25% of familial breast cancer. This study made use of a model, proposed by Antoniou *et al.* (2002), that includes BRCA1 and BRCA2 genes as well as a polygene.

The families used in the study by Antoniou *et al.* (2002) were ascertained through two different methods; (a) cases of breast cancer diagnosed under the age of 55 and (b) families with 2 or more cases of breast cancer, with at least one diagnosed under the age of 50. The maximum likelihood estimation allowed for bias occurring from the use of the two ascertainment methods by (a) maximising the conditional likelihood of observing the disease phenotypes and mutation status given the disease phenotype of the index case, and (b) conditioning the likelihood on the disease phenotypes of all of the family.

The polygenic component was used to attempt to represent the variations, both adverse and beneficial, of alleles in several genes. It was acknowledged that this polygenic component did not explain the remainder of the familial variation but it was thought that it may explain a larger proportion of the variation than the major genes that have already been identified. The polygenic component was assumed to have both adverse and beneficial effects on risks of breast and ovarian cancer.

The premium rates obtained, for BRCA1 and BRCA2 mutations only, agree with results obtained by both Macdonald, Waters & Wekwete (2003a, 2003b) and Gui *et al.* (2006) and the effects of the polygenic component were shown to account for much of the remaining variation.

Macdonald and McIvor found that, for some BRCA1 and BRCA2 mutation carriers, the premium was lower for those with a beneficial polygenotype than for non-mutation carriers with an adverse polygenotype. Beneficial polygenotypes were shown to reduce the premium significantly, and most mutation carriers with this polygenotype were found to be insurable at most ages and for most policy terms. The adverse polygenotype, however, resulted in premiums for carriers being even higher than previously reported.

Through simulation, the premiums for an applicant with a family history were estimated. The premiums, under the full polygenic model, did not increase by much. It was suggested that this was due to the definition of family history, as an increase in the number of affected first degree relatives, in the family history definition, increased the premium rates.

#### **Subramanian *et al.* (2000) and Lemaire *et al.* (2000)**

Both of these studies investigated the effects on life insurance of possessing a family history of breast and ovarian cancer.

Subramanian *et al.* (2000) found that the cost of adverse selection in life insurance, in the case of breast and ovarian cancer, would be minimal if the insurer requested information about the family history of cancer, including the age at onset for all first degree relatives.

Lemaire *et al.* (2000) used onset rates from a study carried out by Claus *et al.* (1994) to estimate premium rates for those with a family history of breast and ovarian cancer. Based on results from epidemiological studies they stated that the key underwriting factors associated with breast and ovarian cancer are: (a) how closely related the affected relatives are; and (b) the ages at onset of the disease. They found that many women with a family history of breast and ovarian cancer can be insured at the standard rate, but that those with two family members with cancer or those with one first degree relative who suffered onset at an early age can only be accepted at substandard rates. They also stated that those with a BRCA mutation will almost always be uninsurable at standard rates. They agreed with Subramanian *et al.* (2000), saying that the insurer should gather as much information as possible about family history in order to avoid adverse selection.

The estimated onset rates used by Lemaire *et al.* (2000) may not be up to date, as genetic testing had not been developed when the study by Claus *et al.* was published in 1994.

### 3.5.2 Huntington's Disease

#### Smith (1998)

Smith investigated life insurance for those with HD. He began by fitting Normal penetrance curves to data from Newcombe (1981). Then, using results from Roos *et al.* (1993), he estimated penetrance and post-onset mortality curves and the resulting extra premiums for life insurance. He concluded that life insurance could still be offered to healthy people who tested positive for HD mutations. He also found that extra mortality associated with HD, as a proportion of normal mortality, increased with increasing term of insurance and decreased with increasing age.

#### Gutiérrez & Macdonald (2004)

As HD is a progressive illness the CI event can be difficult to define. Harper (1996) outlined the progression of HD and categorised it into three stages, see Table 4.1, where stage 1 is defined by manifestations of the disease being present but not debilitating and stage 3 is defined by symptoms so severe that the sufferer is completely dependent on others.

The CI event is usually taken to be at either stage 2 or stage 3 of Harper's progression. The age at which these stages occur is usually taken to occur at about 5 or 10 years after onset respectively. Gutiérrez & Macdonald (2004) interpreted this delay after onset using an accelerated lifetime model, which will be explained in more detail in Section 6.5.

The model of CI used is the same as in Gutiérrez & Macdonald (2003), with the addition of a separate state, onset of HD, with subsequent transition into the 'CI event' state being possible, see Figure 3.2.

Another factor that is taken into account by Gutiérrez & Macdonald (2004) is the number of CAG repeats. The huntingtin mutation occurs because of extra CAG repeats in the huntingtin gene. The number of CAG repeats is known to affect



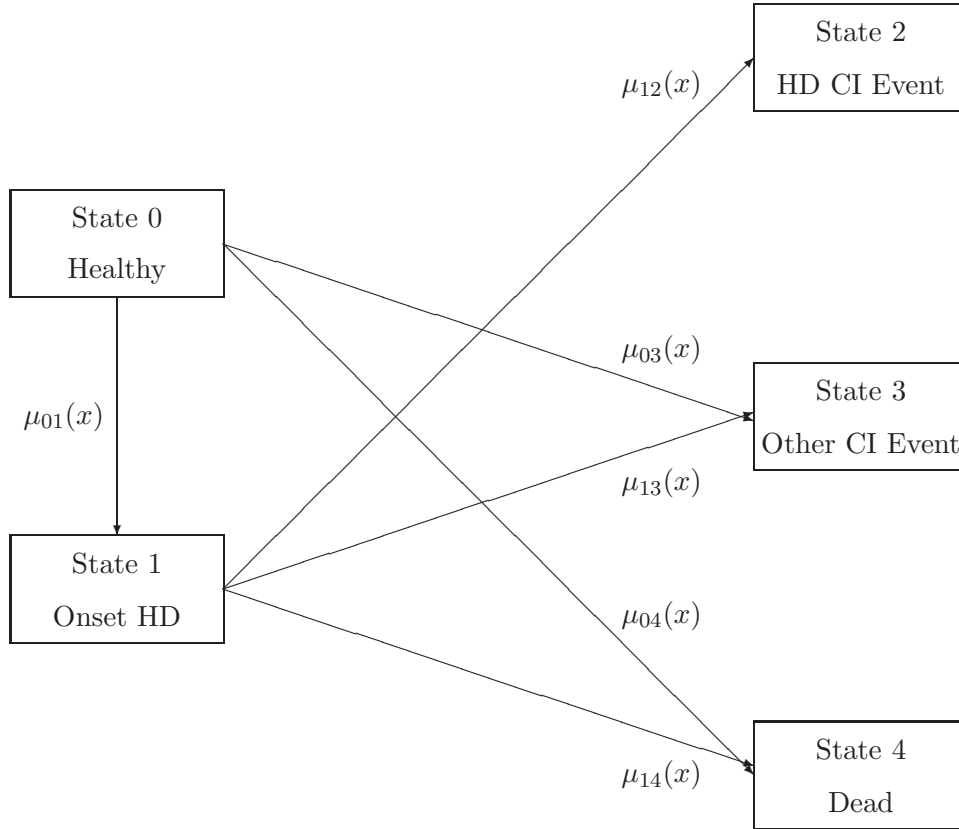


Figure 3.2: A multiple state model for HD in Critical Illness insurance.

the age at onset, an increase in the number of CAG repeats being associated with a decrease in the age at onset (this is explained in more detail in Section 4.1.1). Gutiérrez & Macdonald split the mutation carrier group into categories by the number of CAG repeats.

The rate of onset, as a function of age and CAG repeat length, and the distribution of CAG repeat lengths in the population were modelled using estimates from Brinkman *et al.* (1997). Brinkman *et al.* (1997) carried out sensitivity analysis to determine whether there was any bias introduced by using 2 or more individuals from the same family by randomly selecting 2 individuals from each family and carrying out the analysis based only on these people. They concluded that no obvious bias was introduced. The survival rates after onset of HD that were used were based on estimates by Foroud *et al.* (1999).

The premium rates based on the CI event being stage 3 of Harper's progression showed that applicants would be insurable over a number of categories of age, term and CAG repeat length, although there were a number of categories with a very

high premium of over 1,000% of the OR premium.

The CAG repeat length was found to affect the premium rate significantly; the longer the CAG repeat the higher the premium. Using stage 2 of Harper's progression as the CI event increased the premiums, as compared to stage 3. Basing the calculation of premiums on a family history of HD only, most applicants are insurable, with the exception of the very young and older persons.

Gutiérrez & Macdonald (2004) found that a moratorium on genetic test results would result in premiums increasing very slightly, for both stages of HD progression, as would a moratorium on genetic test results and family history. They noted that the latter would result in increased premiums caused by all high-risk people being part of the standard risk group, not only as a result of adverse selection. Both moratoria affect the small insurance market more than the large market, as found previously.

Gutiérrez & Macdonald (2004) also estimated the costs of adverse selection given the two moratoria, assuming an extreme rate of adverse selection. They found that the premiums increased by as much as 0.1%, given moratoria on genetic tests only, and 0.35% under a moratorium on both genetic tests and family history. The costs were higher assuming an extreme rate of adverse selection than when assuming a moderate rate of adverse selection, but the costs were still found to be relatively small.

The increase under a moratorium on family history was found to be so large because those with a family history can buy insurance at the standard rate. They were assumed to do this at a rate of 0.25 a year, which is high compared with the rate of genetic testing.

All of these papers concluded that the costs of adverse selection were minimal for each disorder; they did note, however, that although the costs were small taking each disorder separately, the costs may add up when including all genetic disorders for which testing will become available.

### 3.5.3 General Discussion

The authors of the papers discussed here were not able to use original data to estimate the required incidence rates, but instead had to derive onset rates and penetrance curves through examination of published tables and plots. The authors often had to use the estimates of the parameters, if given in the study that they used, or to examine Kaplan-Meier plots and fit an appropriate curve to the values obtained from the plots. Variances of parameter estimates are not routinely published in the epidemiological literature and as the studies discussed here were not able to use the original pedigree data they could not estimate the variances of any parameter estimates themselves.

The parameters that are usually provided in published papers typically relate to the distribution of the age at onset. As with all parameter estimation the distribution of the age at onset is dependent on the data available, and as a result different data may give different estimates of the parameters of the distribution. To determine how reliable parameter estimates are their variances, or standard deviations, are often used. This information would be valuable when using the distributions for further work, such as the estimation of premium rates.

Applications to GAIC for the use of a genetic test for insurance purposes requires that there is evidence that the genetic test is reliable in determining increased risk. Having information available about the statistical reliability of the parameter estimates should ideally form part of this evidence.

To determine the variance of the premium rates the variance and covariance of the parameter estimates (the information matrix) is required, in the parametric case. Using bootstrapping methods (Section 5.8) the sampling distribution of the premium rates can be determined using the information matrix of the parameter estimates.

Studies may sometimes publish the variance, or standard deviation, of the parameter estimates but the covariance of these estimates is not normally given. Lu *et al.* (2006) investigated a possible approach to overcoming the lack of information about the covariance of the parameter estimates. By investigating different correlation structures of the parameter estimates they were able to obtain information

about the sampling distribution of the premium rates.

The ideal scenario would be for genetic studies to publish the full covariance matrix of the parameter estimates. This would enable the estimation of the variance of the premium rates. As noted above, this is unfortunately not usually done. The information matrix could be determined from the original data but pedigree data are not normally available to those researching the effects of genetics on insurance. In fact as far as we know no study concerning genetics and insurance has used original pedigree data and resulting pedigree likelihoods to determine the distribution of the age at onset. Gui and Macdonald (2002) used published pedigrees but they used them to calculate Nelson-Aalen estimates of the incidence rates of early-onset Alzheimers disease and discuss what information pedigrees should give.

The work to be undertaken here does use original pedigree data, provided by Jacki Needs at Cardiff University, to estimate the onset rates. Likelihood methods are used to estimate the parameters of the distribution of age at onset (See Section 5.3) and the information matrix. This information matrix can then be used to determine the variance of the premium rates (using bootstrapping methods) and, in turn, provide confidence intervals. We will also be able to show how estimated extra premiums depend on the estimated onset rates and how this relationship changes with age at application and term of policy.

The ability to make some adjustment for ascertainment bias and to gain an impression of the impact on critical illness insurance premiums is, as noted in Section 2.3.2, another major advantage in having data from which to estimate the appropriate parameters. The possible methods of dealing with ascertainment bias, including the method which we use here, are detailed in Section 5.9.

Being able to use original pedigree data, rarely possible in genetics-related insurance studies, will not only enable us to determine a measure of the reliability of the premium rates estimated but will also enable us to see how much of an effect our chosen method of ascertainment bias adjustment has on these results.

# Chapter 4

## Epidemiological Review

Four autosomal dominant genetic diseases are considered here; Huntington's Disease (HD), Breast and Ovarian Cancer (BCOC), Adult Polycystic Kidney Disease (APKD) and Familial Adenomatous Polyposis (FAP). BCOC and HD are both on the list of disorders important to insurers (see Section 3.2). APKD was on this list but was removed due to the availability of ultrasonographic screening for diagnosis, while FAP is no longer being actively considered by the ABI but is still of interest to us. It is of interest to determine how the availability of genetic testing for these disorders may affect premiums.

As discussed in Section 3.3, many epidemiological studies focus on these important genetic diseases. The information obtained from these studies is important for use in clinical settings, genetic counselling and further research. The quantities of most interest are the age at onset of the disease, duration of survival after onset, the penetrance of the disease and the prevalence of mutations.

Here we will review previous research into the four diseases of interest to us, looking in particular at the quantities of interest mentioned above.

### 4.1 Huntington's Disease

Huntington's Disease (HD) is a rare inherited neurological disorder. The possible symptoms of HD range from slurred speech and uncontrolled movements to depres-

sion and immobility. As a result HD can be difficult to diagnose early as many of these symptoms are common to other illnesses.

HD is an autosomal dominant disorder: it is passed from the parents to children through what has become known as the huntingtin gene. Every person has two copies of the huntingtin gene, one inherited from each parent. If there is a HD mutation in either of the two copies, that copy produces a faulty version of the huntingtin protein. It is this faulty protein that causes the deterioration of the brain over time, specifically the control of motor and mental processes.

When a patient starts to suffer the symptoms of HD they are said to have suffered onset of the disease, although there is the possibility of delayed diagnosis. Table 4.1 shows the progression of HD in three stages as defined by Harper (1996). In terms of critical illness insurance stages 2 or 3 can be thought of as the insurable event, depending on policy conditions. That is, it is a convenient proxy.

#### 4.1.1 Genetic Testing and Penetrance

The Huntington mutation was identified on chromosome 4 (Gusella *et al.*, 1983), and later found to be specifically at locus 4p16.3 (The Huntington’s Disease Collaborative Research Group (HDCRG), 1993).

The faulty huntingtin protein is formed as a result of a trinucleotide expansion in the Huntington gene. This means that extra ‘words’ are inserted in the gene, in this case the extra word being the trinucleotide CAG. The CAG trinucleotide can be repeated a number of times at a specific location in the gene, and it has been shown that the number of CAG repeats affects the severity of the symptoms of HD (Andrew *et al.*, 1993, Illarioshkin *et al.*, 1994, Furtado *et al.*, 1996).

Most studies agree that those with the normal phenotype (unaffected) have fewer than 36 CAG repeats (Stine *et al.*, 1993, Masuda *et al.*, 1995, Brinkman *et al.*, 1997, Alonso *et al.*, 1997, Jakab *et al.*, 1999 and Langbehn *et al.*, 2004) although some have disagreed. Barron *et al.* (1993) found a range of 35 – 62 CAG repeats for affected subjects and Snell *et al.* (1993) found a range of 30 – 70.

HD is fully penetrant (Roos *et al.*, 1991), however, Rubinsztein *et al.* (1996) found

Table 4.1: Stages of progression of Huntington’s Disease for a typical patient.  
Source: Harper (1996). (Note: Chorea refers to brief, repetitive, jerky, or dancelike uncontrolled movements caused by muscle contractions.)

Stage	Clinical Features
1	Presentation with initial neurological or psychiatric symptoms Main features remain similar to those at presentation Chorea more prominent than other motor abnormalities Patient largely independent for most activities Burden on family mainly result of psychiatric problems Death rare except for suicide
2	Motor disorders more generalised Physical disability becomes major Patient dependent on others for many activities Burden on family both physical and psychological Death often from unrelated causes
3	Severe generalised motor disorder Physical disability severe to total Patient completely dependent for all aspects of life Burden on carers mainly physical Death frequent at any point

that HD was not fully penetrant for those with 36 – 39 CAG repeats. Brinkman *et al.* (1997) found reduced penetrance for those with fewer than 42 repeats, while McNeil *et al.* (1997) found reduced penetrance for 37 – 39 repeats.

The U.S. Venezuela Collaborative Research Project (2004) found that persons with fewer than 34 CAG repeats do not suffer symptoms, persons with 35 – 39 repeats have incomplete penetrance of HD and those with  $\geq 40$  CAG repeats show full penetrance.

The different conclusions reached by these studies could possibly be caused by different laboratory procedures for estimating CAG repeat length (Brinkman *et al.*, 1997). Rubinsztein *et al.* (1996) stated that the estimates made in previous studies may not be reliable as the methods used to determine number of CAG repeats made assumptions that were not valid.

The genetic test for the huntingtin mutation is ‘targeted mutation analysis’. This is used because the huntingtin gene has been identified and there is no other cause of HD, and mutations have a 100% detection rate. If a negative result is received then the person will not develop HD, while a positive result, for a large enough number of CAG repeats indicates that the person will develop HD if they live long enough.

The problem with HD is that although the test is accurate and reliable, a person may be reluctant to know that they have the mutation as there is no cure. The actual proportion of those at risk of HD that undergo testing is about 5 – 24% (Craufurd *et al.*, 1989, Tyler *et al.*, 1992 and Maat-Kievit *et al.*, 2000, Creighton *et al.*, 2003). Creighton *et al.* (2003) found that the uptake for the prenatal test is approximately 2%.

Binedell *et al.* (1998) looked at the differences between two groups, test requesters and non-requesters drawn from a HD register. They found that they differed in: their knowledge of the availability of testing (people who knew more about testing availability were more likely to take the test), their perceived attitudes of family members to testing, their perceived stressfulness of being at risk and how they thought they would cope with the results (non-requesters were more likely to say that they could not cope). Taylor (2004) also investigated the decision-making process undergone by people considering presymptomatic genetic testing for HD.



She found that a person's decision about whether to take the test or not depended on their attitudes to a number of different factors including feelings about family loyalty and feelings about the timing of the test.

### 4.1.2 Onset Rates

Much research has been undertaken to estimate the distribution of the age at onset of HD for persons from HD families. Table 4.2 summarises a number of these studies, in terms of the number of subjects, the mean age at onset and the standard deviation of the age at onset. The mean age at onset ranges from 33.8 to 43.97 years. This variation may be because of various methods of analysis, as well as geographical and temporal differences.

Newcombe (1981) stated that ascertainment bias was introduced in a number of studies because they only used persons who had suffered onset, from the affected families sampled, to estimate the age at onset. They did not include people who had not suffered onset. He estimated the age at onset while attempting to remove this bias. He found the mean age at onset to be 45.96 years, including all available people and 48.86 years, when probands (those through which the pedigree was identified) were removed. These mean ages are higher than those obtained from other studies, in particular they are 4.6 – 7.5 years greater than previous work that did not attempt to remove this bias.

Myers *et al.* (1985) found that about a quarter of those affected by HD suffered onset after age 50 and that the average age at onset for late onset sufferers was 57.5.

### CAG Repeats

Table 4.3 shows the correlation between age at onset and number of CAG repeats found in some studies. It is clear that higher numbers of CAG repeats are associated with earlier ages at onset. Brinkman *et al.* (1997) also found that the relationship between  $\log(\text{age at onset})$  and  $\log(\text{CAG repeat length})$  is strong, (correlation between them is  $-0.73$ ).

Although the relationship is usually strong there has been some disagreement about

Table 4.2: Mean age at onset of HD in published studies.  $N$  is the number of subjects in the study,  $\mu$  is the mean age estimated by that study and  $\sigma$  is the standard deviation. Source: Gutiérrez & Macdonald (2002a)

<b>Reference</b>	$N$	$\mu$	$\sigma$
Adams <i>et al.</i> (1988)	611	38.66	11.69
Andrew <i>et al.</i> (1993)	360	41.5	12.4
Bell (1934)	460	35.51	12.38
Bolt (1970)	265	42.7	13.2
Brackenridge (1971)	344	33.8	
Brinkman <i>et al.</i> (1997)	728	41.0	13.0
Brothers (1964)	206	37.2	
Dewhurst <i>et al.</i> (1970)	102	39.0	
Farrer & Conneally (1985)	569	38.0	11.1
Feigin <i>et al.</i> (1995)	129	36.7	12.9
Folstein <i>et al.</i> (1987)	217	40.25	12.9
Foroud <i>et al.</i> (1999)	2,068	40.0	12.0
Marder <i>et al.</i> (2000)	960	40.8	12.4
Morrison <i>et al.</i> (1995)	143	43.6	13.5
Myers <i>et al.</i> (1985)	243	40.9	
Panse (1942)	446	36.2	12.3
Reed <i>et al.</i> (1958)	262	35.3	9.8
Roos <i>et al.</i> (1991)	1,020	39.5	12.1
Stevens (1977)	162	43.43	10.26
Venters (1971)	123	38.8	10.11
Walker <i>et al.</i> (1981)	204	41.2	12.7
Wallace (1972)	144	39.5	
Wendt <i>et al.</i> (1959)	762	43.97	10.9
Wendt & Drohm (1972)	802	43.39	10.08

Table 4.3: Correlation between CAG repeat length and age at onset of Huntington’s Disease in published studies. The Kehoe study published the results separately for those that inherited the mutation from their mother (maternal transmission) and father (paternal transmission).

Reference	Correlation ( $\rho$ )
Andrew <i>et al.</i> (1993)	−0.7
Brandt <i>et al.</i> (1996)	−0.72
Kehoe <i>et al.</i> (1999)	−0.771 (maternal transmission)
Kehoe <i>et al.</i> (1999)	−0.751 (paternal transmission)
Kiebertz <i>et al.</i> (1994)	−0.82
Myers <i>et al.</i> (2004)	−0.81
Simpson <i>et al.</i> (1993)	−0.58
Snell <i>et al.</i> (1993)	−0.74
Stine <i>et al.</i> (1993)	−0.65

whether this enables the prediction of age at onset based on the number of CAG repeats. Brinkman *et al.* (1997) suggest that prediction based on CAG repeat length alone is reliable and Penney *et al.* (1997) found that CAG repeat length accounts for 78% of the variation in age at onset, while Andrew *et al.* (1993) states that the number of CAG repeats accounts for less than 50% of the variation and is therefore not reliable enough to be used for prediction.

Langbehn *et al.* (2004) developed a parametric model to predict age at onset given the number of CAG repeats. The model suggested that there is a high probability that a person with fewer than 40 CAG repeats will not suffer onset in their lifetime, however long.

### Sex of Affected Parent

The sex of the affected parent has been found to affect the age at onset of their offspring, Jones & Phillips (1970) found that when all cases of onset at 21 years or younger were removed there was no longer an effect of the sex of the transmitting parent. Stevens (1977) found that there was also an effect of the sex of the affected

grandparent and that if the mutation was transmitted from grandfather to father to offspring then the age at onset would be much earlier; Newcombe *et al.* (1981) agreed with these findings. Krawczak *et al.* (1991) found that the affected parent's age at onset and sex were significantly correlated with the age at onset of the offspring.

## 4.2 Breast and Ovarian Cancer

Breast and ovarian cancer are serious diseases that affect women, though there are also rare instances of male breast cancer. These cancers, as with most others, can be fatal if not diagnosed early enough. Many investigations have been carried out to determine what may cause breast and ovarian cancer.

The pedigrees of breast and ovarian cancer families suggest three possible patterns of disease; sporadic, familial and hereditary.

A sporadic pattern occurs when there is no regular pattern of cancer in the family. A familial pattern indicates there is family history of cancer but does not suggest a Mendelian pattern of disease while a hereditary pattern suggests that there is a Mendelian pattern of disease.

The existence of a family history of breast and ovarian cancer has been shown to affect a person's risk of developing breast and ovarian cancer.

Anderson *et al.* (1992) estimated the penetrance of breast cancer among first degree relatives of breast cancer patients as 11.6%. Sutcliffe *et al.* (2000) found that the penetrances of breast and ovarian cancer for people with more than 2 cases of ovarian cancer in the family were 15% and 11%, respectively.

Claus *et al.* (1993) estimated the penetrance of developing breast cancer for a woman with 1 or 2 first degree relatives with ovarian cancer as 13% and 31%, respectively. They also found that a woman has a 40% risk of breast cancer by age 79 if she has 1 first degree relative with ovarian cancer and 1 with breast cancer, if the relatives suffered onset in their 30's.

The Collaborative Group on Hormonal Factors in Breast Cancer (2001) have shown

that the risk of developing breast cancer increases as the number of relatives with the disease increases. They found that a person with 1 affected first-degree relative was 1.8 times more likely to develop breast cancer than a person with no affected first-degree relatives. This ratio increased as the number of affected first-degree relatives increased, to 2.89 and 3.9 for 2 and 3 or more affected relatives, respectively.

Slattery and Kerber (1993) found that the risk of breast cancer decreased the more distant the relative with breast cancer was. The odds ratio was 2.45 for women with a first degree relative with BC, 1.28 for women whose nearest affected relative was second degree and 1.35 when the nearest affected relative was third degree.

Through investigation of familial breast and ovarian cancer a number of studies found evidence to suggest that there is an autosomal dominant allele that leads to increased susceptibility to breast and ovarian cancer (Claus *et al.*, 1991 & 1994; Piver *et al.*, 1993). These studies and others like them led to the search for breast and ovarian cancer-causing genes.

#### **4.2.1 Genetic Testing**

A study using a large number of families with cases of early onset breast cancer assigned the BRCA1 mutation, through linkage analysis, to chromosome 17 (Hall *et al.*, 1990). It was subsequently found to be at position 17q21 (Miki *et al.*, 1994). Although this mutation accounted for a high proportion of hereditary breast and ovarian cancer, it was found that families with a high incidence of male breast cancer did not carry this mutation (Stratton *et al.*, 1994). This launched a further search for other mutations which led to the discovery of a second mutation, BRCA2. It was linked to chromosome 13 in 1994 (Wooster *et al.*, 1994) and isolated one year later at position 13q12.3 (Wooster *et al.*, 1995).

The BRCA1 and BRCA2 mutations only account for a small proportion of all breast and ovarian cancers. According to the Collaborative Group on Hormonal Factors in Breast Cancer (2001), 7 – 10% of ovarian cancers (OC) and 5 – 10% of breast cancers (BC) are due to BRCA1 and BRCA2 mutations.

There are many known mutations in BRCA1 and BRCA2 and some work has been

carried out to determine if there are mutations that are unique to specific populations. It was found that there were three types of BRCA1/2 mutations that were associated with Ashkenazi Jews, 185delAG, 5382insC on BRCA1 and 6174delT on BRCA2 (Struewing *et al.*, 1995, Tonin *et al.*, 1995, Neuhausen *et al.*, 1996) and that if a Jewish woman does not carry one of these mutations then it is unlikely that she will carry a different BRCA1/2 mutation (Narod & Foulkes, 2004).

There are two main types of genetic tests; mutation analysis and sequence analysis (Petrucelli *et al.*, 2005). Mutation analysis tests for the three Ashkenazi founder mutations, while sequence analysis tests for BRCA1 and BRCA2 sequence alterations as well as five specific large genomic BRCA1 rearrangements. Another available test is the protein truncation test (PTT), developed by Hogervorst *et al.* (1995), which, although inexpensive and quick, has been restricted to screening large exons of BRCA1 and BRCA2 by many clinics (Narod & Foulkes, 2004).

Mutation analysis has a detection rate of 90% (Frank *et al.*, 1998), while sequence analysis has a detection rate of more than 88% in families shown to have a linkage to BRCA1 or BRCA2 (Petrucelli *et al.*, 2005). If a negative result is obtained by mutation analysis for a person whose family is known to carry a mutation then the person's risk of developing breast or ovarian cancer is reduced but is still at least the same as the risk of the general population. If the family is not known to carry a mutation, sequence analysis may be recommended depending on the person's family history and their risk of having a mutation. If a positive result is obtained a person's risk of breast and ovarian cancer is greatly increased.

### 4.2.2 Prevalence

BRCA1 and BRCA2 are rare; the frequencies of BRCA1 and BRCA2 mutations are estimated to be 0.051% and 0.068% respectively, (Antoniou *et al.*, 2002). Antoniou *et al.* (2000) found the frequencies to be slightly higher, 0.128% and 0.172%, respectively. They are found to be more common among families with histories of breast and ovarian cancer.

In a study carried out using families with at least 4 cases of breast cancer, Ford *et al.* (1998) found that 52% of familial breast cancer was due to BRCA1 and 35% was

due to BRCA2. Therefore 13% was due to something else, perhaps another gene or combination of genes. When using only families with four or five cases of breast cancer, male or female, these proportions changed, 55% due to BRCA1 and 12% due to BRCA2. When they considered only families with four or five cases of breast cancer, female only, they found that only 32% were due to BRCA1 and 9% were due to BRCA2. The Anglican Breast Cancer Study Group (2000) found that only 17% of breast cancer familial relative risk was attributable to BRCA1 and BRCA2.

A number of studies have estimated the proportion of people with breast and ovarian cancer that carry either BRCA1 or BRCA2 mutations. One study stated that a BRCA1 mutation was found in 16% of women with breast cancer, and in 14% of women who suffered breast cancer before age 40. The mutation was also found in 40% of families with breast and ovarian cancer (Couch *et al.*, 1997). Liede *et al.* (2002) found that 6.7% of women with breast cancer and 15.8% of women with ovarian cancer had one of the mutations (BRCA1 or BRCA2).

Malone *et al.* (1998) found a BRCA1 mutation in 6.2% of those diagnosed with breast cancer before age 35 and in 7.2% of those diagnosed with breast cancer before age 45, who had a first degree family history. A more recent study found mutations in 9.4% of people diagnosed with breast cancer before age 35 (3.4% in BRCA2 and 5.9% in BRCA1) and in 12.0% of those diagnosed before age 45 who also had a first degree family history (4.9% in BRCA2 and 7.1% in BRCA1) (Malone *et al.*, 2000).

Another study found mutations (BRCA1 and BRCA2) in 5.9% of women with diagnosis of breast cancer before age 36, and in 4.1% of women diagnosed between ages 36 and 45. They also found a mutation in 11% of patients with a first degree relative who was diagnosed with breast or ovarian cancer before age 60 and in 45% of patients with 2 or more affected first or second degree relatives (Peto *et al.*, 1999).

The prevalences of the Ashkenazi founder mutations among Ashkenazi Jewish women have been estimated to be 1.09%, 0.13% and 1.52% for the mutations 185delAg, 5382insC and 6174delT, respectively (Roa *et al.*, 1996). Other studies estimated similar prevalences, 1% for 6174delT, (Oddoux *et al.*, 1996) and 0.9% for 185delAG, (Struwing *et al.*, 1995).

It was also found that over 2% of Ashkenazi Jewish women carry a mutation in BRCA1 or BRCA2 (Struewing *et al.*, 1997) and that the mutations are present in about 20% of Ashkenazi Jewish women with early onset breast cancer (Offit *et al.*, 1996).

With the exception of the studies concerning Ashkenazi Jewish women the estimates of the proportion of people with breast or ovarian cancer who carry either the BRCA1 or BRCA2 mutations are very similar. These values give a measure of the prevalence of BRCA1 and BRCA2 amongst breast and ovarian cancer sufferers.

### 4.2.3 Penetrance

Carrying a BRCA1 or BRCA2 mutation confers a very high probability of developing breast or ovarian cancer. Table 4.4 shows penetrance estimates for breast cancer, given the presence of BRCA1 or BRCA2 mutations, from a number of studies, and Table 4.5 shows the corresponding results for ovarian cancer. In both tables the estimates in the right-hand column were obtained from studies which did not differentiate between BRCA1 and BRCA2.

Another study carried out by Antoniou *et al.* (2005) estimated the penetrance for the Ashkenazi founder mutations using the data from their 2003 study. They found that the penetrance of breast cancer by age 70 was 64% for 185delAG, 67% for 5382insC and 43% for 6174delT. The penetrance of ovarian cancer was 14%, 33% and 20% for 185delAG, 5382insC and 6174delT, respectively, while Struewing *et al.* (1997) found that the penetrance of breast and ovarian cancer for those with the Ashkenazi founder mutations was 56% and 16%, respectively.

These penetrance estimates vary widely and it is difficult to determine which is the most reliable. However, comparing all of them with the population lifetime risks of about 11% for breast cancer and 2% for ovarian cancer it is clear that the probability of developing either cancer is greatly increased if a mutation in either gene is present. In fact Claus *et al.* (1998) investigated the differences in penetrance for both mutation carriers and non-mutation carriers and found that the penetrance of breast cancer by age 80 was 63% and 9%, respectively.



Table 4.4: Breast cancer penetrance in published studies. The penetrance is defined at age 70 for all studies except for Claus *et al.* (1998), Risch *et al.* (2001) and Anglican Breast Cancer Study Group (2000) where it is defined at age 80.

Reference	BRCA1	BRCA2	BRCA1 or BRCA2
Lalloo <i>et al.</i> (2006)	84%	91%	
Antoniou <i>et al.</i> (2006)	72%	75%	
Marroni <i>et al.</i> (2004)	39%	44%	
Antoniou <i>et al.</i> (2003)	65%	45%	
Antoniou <i>et al.</i> (2002)	35.3%	50.3%	
Risch <i>et al.</i> (2001)	68%	13.9%	
Anglican Breast Cancer SG (2000)	48%	74%	
Antoniou <i>et al.</i> (2000)	45%		
Ford <i>et al.</i> (1998)			84%
Claus <i>et al.</i> (1998)			63%
Easton <i>et al.</i> (1995)	85%		
Narod <i>et al.</i> (1995)	71%		

#### 4.2.4 Treatment

Surgery may be advised when breast or ovarian cancer is diagnosed, but it can also be advised as a preventative treatment for mutation carriers. Prophylactic surgery has been shown to reduce a carrier’s risk of developing breast and ovarian cancer. Rebbeck *et al.* (2004) found that having a prophylactic oophorectomy and mastectomy reduces the risk of breast cancer by as much as 95%; they also noted that having a mastectomy alone reduces the risk of breast cancer by 90%. Rebbeck *et al.* (2002) found that an oophorectomy reduces the risk of breast cancer by approximately 50% and ovarian cancer by about 90%.

Schrag *et al.* (1997) stated that a 30-year old woman with a BRCA1 or BRCA2 mutation increases their life expectancy by between 2.9 and 5.3 years if they undergo a prophylactic mastectomy and between 0.3 and 1.7 years if they undergo a prophylactic oophorectomy. These gains were shown to decline with age and to be

Table 4.5: Ovarian cancer penetrance in published studies. The penetrance is defined at age 70 for all studies except for Claus *et al.* (1998), Risch *et al.* (2001) and Anglican Breast Cancer Study Group (2000) where it is defined at age 80.

Reference	BRCA1	BRCA2	BRCA1 or BRCA2
Lalloo <i>et al.</i> (2006)	60%	26%	
Antoniou <i>et al.</i> (2006)	38%	49%	
Marroni <i>et al.</i> (2004)	43%	15%	
Antoniou <i>et al.</i> (2003)	39%	11%	
Antoniou <i>et al.</i> (2002)	25.9%	9.1%	
Risch <i>et al.</i> (2001)	36%	9.7%	
Anglican Breast Cancer SG (2000)			22%
Antoniou <i>et al.</i> (2000)	66%		
Ford <i>et al.</i> (1998)			27%
Easton <i>et al.</i> (1995)	63.3%		
Narod <i>et al.</i> (1995)	42%		

minimal at age 60. Armstrong *et al.* (2004) found that a prophylactic oophorectomy lengthened the life expectancy by between 3.34 and 4.65 years depending on age at surgery. Rebbeck *et al.* (1999) agreed, showing that a bilateral prophylactic oophorectomy reduced the risk of breast cancer by about 50%. The risk of breast cancer was only slightly reduced for those who underwent surgery after age 50.

Some investigation has been carried out to determine whether surveillance (ovarian screening and mammography) is reliable. Meeuwissen *et al.* (2005) discovered that ovarian cancer surveillance has limited sensitivity and gives a high number of false positive findings. They recommended that a prophylactic oophorectomy is the optimal risk-reducing strategy for women at high risk of breast and ovarian cancer. Vasen *et al.* (2005) also found that surveillance for ovarian cancer was not effective as all of the screen-detected tumours were advanced cancers. They did, however, find that mammography was quite a reliable screening method for breast cancer as only 20% of tumours detected had lymph node metastasis.

Lynch *et al.* (2006) investigated the attitudes to genetic testing and prophylactic surgery of people from hereditary breast and ovarian cancer families with BRCA mutations. They found that before testing 23% had a prophylactic bilateral mastectomy or oophorectomy or both, and of these 53% were subsequently found not to have either mutation. After testing, 52.9% of mutation carriers and 0% of non-carriers underwent prophylactic surgery.

#### 4.2.5 Other Genes

As BRCA1 and BRCA2 account for only a small proportion of breast and ovarian cancer it is clear that there may be other breast cancer susceptibility genes that are, as yet, undiscovered. It is, however, unclear whether they are rare mutations which confer a high risk or common low risk mutations.

It has been suggested that breast cancer families who do not have BRCA1 or BRCA2 mutations may carry mutations which act in a more complex manner. Peto (2002) suggests that these genes may influence breast cancer susceptibility through gene-gene or gene-environment interactions.

Antoniou *et al.* (2002) found that the best model for familial breast cancer was one that incorporated BRCA1 and BRCA2 mutations as well as a polygenic component, where the polygenic component affected BC but not OC. The polygenic component represented a large number of low risk polymorphisms acting multiplicatively. They also found that the presence of a third major breast cancer susceptibility gene (BRCA3) was unlikely.

In Section 3.5.1 we discussed a paper investigating insurance costs based on the assumption that there is a polygenic component in the model of breast cancer.

### 4.3 Familial Adenomatous Polyposis

Familial adenomatous polyposis (FAP), or familial polyposis coli (FPC), is a genetic disorder in which polyps develop in the colon and rectum. There is a high risk that, if nothing is done, the polyps will progress to colorectal cancer (CRC); the

lifetime risk of colon cancer is nearly 100% (Järvinen, 2004). Other, less common, manifestations of FAP include; osteomas, epidermoid cysts, retinal pigmentation, gastric and duodenal polyps, desmoid tumors, brain tumors, periampullary adrenal and thyroid carcinomas (Arvantis *et al.*, 1990). Colorectal cancer is, however, the main cause of death in people with FAP and is defined as the CI event for FAP.

Järvinen (1992) found that 0.53% of all colorectal carcinoma was associated with FAP, while Bülow (2003) estimated this proportion to be 0.07%. The prevalence of FAP, based on national registries, has been determined by a number of studies whose estimates lie in the range 2.29 – 4.65 per 100,000 (Burn *et al.*, 1991, Bülow 1996, Järvinen 1992 and Bülow, 2003). Previous studies have provided a range of estimates of the incidence of FAP, 0.62 – 2.38 per million (Järvinen, 1992 and Bülow, 2003).

It has been shown that persons with affected first-degree relatives have an increased risk of developing colon cancer. Those who have a first-degree relative who has adenomatous polyps have an increased risk of colon cancer of 72, 74 or 78% (Fuchs *et al.*, 1994, Ahsan, 1998 and Winawer, 1996, respectively) while those who have a first-degree relative who has had colon cancer have an increased risk of colon cancer of approximately 218, 225 or 326% (Bonelli *et al.*, 1988, Winawer, 1996 and Ponz de Leon *et al.*, 1987, respectively)

It has also been shown that as the number of affected first-degree relatives a person has increases, the risk of colon cancer increases. St John *et al.* (1993) estimated the odds ratio for colon cancer risk to be 1.8 given 1 first-degree relative and 5.7 given 2 first-degree relatives, compared to controls. Fuchs *et al.* (1994) estimated these odds ratios as 1.72 and 2.75, respectively, while Winawer *et al.* (1996) found them to be 1.78 and 3.25, respectively.

Mitchell *et al.* (2005) found that 9.4% of those with colorectal cancer had a first-degree relative with colorectal cancer, while 28.8% had an affected first- or second-degree relative.

A patient is diagnosed with FAP if they have more than 100 colorectal adenomatous polyps (detected by sigmoidoscopy or colonoscopy), or have fewer than 100 polyps and a relative with FAP. FAP is usually diagnosed using screening methods; how-

ever, the availability of genetic testing brings a new dimension to the surveillance and diagnosis of FAP.

### 4.3.1 Genetic Testing

The autosomal dominant gene associated with FAP has become known as the APC (adenomatous polyposis coli) gene. It was linked to chromosome 5 in 1987 (Bodmer *et al.*, 1987), and subsequently found to be at location 5q21, (Groden *et al.* 1991, Kinzler *et al.*, 1991 and Nishisho *et al.*, 1991). Until recently it was believed that all cases of FAP were caused by the inheritance of mutations in the tumour-suppressor APC gene (Chung, 2000). Now however, there is some evidence that points to a proportion of FAP being due to autosomal recessive inheritance of mutations in MUTYH, a DNA repair gene.

There are many mutations in the APC gene. Most of the mutations produce a truncated APC protein through point substitutions or frameshift mutations (Armstrong *et al.*, 1997). A person with FAP may have inherited one of these many mutations in the APC gene or may have a *de novo* gene. The rate of new mutations is estimated as between 4 and 9 mutations per million gametes per generation (Bisgaard *et al.*, 1994; Björk *et al.*, 1999). The same studies provided estimates of the proportion of FAP cases that have a new mutation ranging from 11 to 25%.

Bertario *et al.* (2003) said that the site of the APC mutation is related to the phenotype of FAP; in fact analysis of this relationship identifies subgroups of FAP patients at high risk for major extracolonic disease.

Burt (2000) listed the genetic tests available for FAP as follows: protein truncation testing, gene sequencing, known mutation in a family and linkage testing. The genetic tests do not have 100% detection rate, in fact, Järvinen (2003) said that the detection rate is 60 – 95%, depending on the method used. Powell *et al.* (1993) said that, given the technology available in 1993, APC mutations were detected in 80 – 90% of FAP families.

Burt (2000) said that once it is known which APC mutation a person has their family members can be tested for this mutation with sufficient accuracy. They can

then be directed to appropriate screening and sometimes prophylactic surgery. If, however, a mutation is not found in the proband then all family members should undergo screening. It has been noted that a positive genetic test result should be followed up by an annual sigmoidoscopy and that a negative result should be followed up by a flexible sigmoidoscopy every 7 – 10 years, to account for any potential error in the genetic testing (Powell, 1993).

### 4.3.2 Diagnosis of FAP

The two main methods of detection used for diagnosis of FAP, sigmoidoscopy and colonoscopy, involve the examination of the colon and the large intestine. Although family screening is recommended for those at 50% risk it has been shown that, apart from those on FAP registers, the majority of patients are symptomatic at time of diagnosis and that FAP is diagnosed based on clinical symptoms in 84% of sufferers (Croner *et al.*, 2005). They found that the commonest presenting features of FAP are alteration of bowel habit (42%) and rectal bleeding (68%), but they stated that both of these symptoms are found in many other conditions and therefore an endoscopy is required to confirm diagnosis.

Family screening is recommended for those at 50% risk of having an APC mutation. Family screening has been shown to be effective in diagnosing FAP earlier. Järvinen (1992) found that the proportion of people with FAP diagnosed by family screening has increased from 10%, between 1966 and 1970, to 56%, between 1986 and 1990. Bernstein & Bülow (2005) found that the prevalence of colorectal cancer is reduced by 55% by family screening.

In most studies the sample is split into two groups, probands and call-up patients. The probands (or propositi) are those that are diagnosed through symptoms, while the call-up patients are diagnosed through screening as they are known to be at 50% risk.

Bisgaard (1994) suggests that polyps nearly always develop by age 40, but Evans (1993) suggested that without a specific genetic test, using genetic markers rather than DNA, screening should be carried out well beyond the age of 40 as there is evidence of some cases of late onset. Vasen *et al.* (1990) found that by age 40, 90%

of the call-up FAP patients that were found to have a mutation, were diagnosed. They suggested that screening should start at age 10 – 12 and continue until age 50, or age 60 for those with a family history of late onset.

Petersen *et al.* (1991) found that the probability of finding polyps in people who had the FAP mutation was 15% at age 10, 82% at age 15, 94% at age 20 and 98% at age 30 and that the probability a person at 50% risk has FAP after a negative sigmoidoscopy was 49% at age 10, 20% at age 20 and 8% at age 30. They found that, combined with linkage marker data, the 50% risk for relatives can be reduced to below 0.5% by age 30 if there is a negative result on sigmoidoscopy and a negative result from linkage analysis.

### **Age at Diagnosis**

Vasen *et al.* (1990) estimated that the mean age at diagnosis of FAP was 24 years for call-up patients and 35 years for probands, while, Heiskanen *et al.* (2000) estimated these mean ages to be 22.8 and 36.8, respectively.

Petersen *et al.* (1991) estimated that the mean age at onset of polyps is 15.9. Bülow (1986) found that the median age at development of polyps, for the call-up patients, was 16, while the median age at first bowel symptom, for the probands patients, was 29. The median age at diagnosis of FAP was age 33, for the probands group, and 19 for the call-up group and the median age at diagnosis of CRC in the two groups was at age 36 and 45.5, respectively.

This indicates that those who undergo family screening are diagnosed with FAP earlier than probands, and that if CRC develops it develops later, due to early surgery to remove pre-cancerous polyps.

### **4.3.3 Colorectal Cancer**

Without surgery a person with FAP has a high risk of colorectal cancer (CRC); 7% affected by age 21, 87% by age 45 and 93% by age 50 (Bussey 1975), with average age 39 at diagnosis.

Table 4.6 shows the proportions of people in the two groups that already had colorectal cancer when they were diagnosed as having FAP. The proportion was very

Table 4.6: Proportion of people with colorectal cancer at time of diagnosis, in the call-up group and in the propositi group.

Reference	Call-up	Propositi
Järvinen (1992)	6.6%	65.5%
Bülow (1986)	3%	69%
Vasen <i>et al.</i> (1990)	4%	47%
Heiskanen <i>et al.</i> (2000)	4.2%	61%
Bülow (2003)	3%	67%

similar for all of the studies shown. These suggest that if diagnosed with FAP after the development of symptoms, the risk of developing colorectal cancer is greater.

Bülow (2003) said that the prevalence of colorectal cancer in Denmark has decreased and the prognosis has improved since the establishment of a Danish polyposis register, and suggests that more registers be set up in order to attempt to eliminate colorectal cancer in persons with FAP.

#### 4.3.4 Other Manifestations of FAP

Congenital hypertrophy of the retinal pigment epithelium (CHRPE) is a symptom associated with FAP. Burn *et al.* (1991) found that in a group of controls, 98% had 0, 1 or 2 small lesions in the eye, while in a group of FAP sufferers 7.6% had 0, 1, or 2 lesions, 4.5% had 3 lesions and 87.8% had 4 or more. They concluded that this could be used as a diagnostic tool for FAP; fewer than 3 lesions suggests a 1 in 13 carrier probability, more than 3 lesions suggests a 2 in 3 carrier probability, and 3 lesions is diagnostic. Tiret *et al.* (1997) agreed that CHRPE could be used as a diagnostic tool. Their criteria were slightly different, however. They said that those with at least 4 lesions or at least 2 lesions, of which one is large, were diagnosed as having FAP. Their criterion had high sensitivity, 0.68, and maximal specificity of 1.

Bülow *et al.* (1995) found that 64% of those with FAP had histologically verified duodenal adenomas, at a median age of 37, while, Bülow *et al.* (2004) found that



65% had duodenal adenomas, at a median age of 38. The high incidence and the fact that the severity of duodenal adenomatosis increased with age prompted them to suggest that prophylactic gastrointestinal endoscopy was justified.

Clark & Phillips (1996) found that desmoid tumours (fibrous tissue tumours which do not spread around the body but can invade nearby tissues) occur in 10% of individuals affected with FAP, while Bertario *et al.* (2001) estimated that desmoids developed in 11.9% of all FAP patients, and that the cumulative rate of desmoids is 20.6% at age 60. They found that females are at a higher risk of developing desmoids than males (odds ratio 2.1), they also found that a family history of desmoids, (8.75), a family history of osteomas, (2.9), and a family history of epidermoid cysts (1.8) were significantly associated with occurrence of desmoids.

These other manifestations of FAP can also be used for diagnosis of FAP, even though they are less prevalent in sufferers of FAP.

#### **4.3.5 Treatment**

Often patients with FAP have prophylactic surgery in the form of a colectomy (removal of the colon), although other forms of surgery are available. Colectomy is often performed in conjunction with other treatments; ileorectal anastomosis (IRA) and ileal pouch-anal anastomosis (IPAA).

The IRA surgery involves the removal of the colon but all or most of the rectum is left in place, with the small intestine being attached to the upper portion of the rectum. IPAA involves the removal of the entire colon and most of the rectum with an artificial rectum, a pouch, being made out of the lower end of the small intestine. The IRA surgery will also involve life-long surveillance of the rectal stump as there is still a risk of cancer of the rectal stump of up to 15% within 25 years (Bülow *et al.*, 2000).

Alarcon *et al.* (1999) found that endoscopic eradication is an appropriate initial treatment for histologically advanced noncancerous neoplasms, or for patients who are not surgical candidates. Pancreas-sparing duodenectomy may be the treatment of choice for patients with carcinoma and those who have failed endoscopic therapy.

The national polyp study found a 76 – 90% decrease in colon cancer incidence after polypectomy (Winawer, 1996)

Järvinen (1985) said that the ideal treatment for FAP was prophylactic colectomy and that it should be performed by age 20 – 25. Bülow *et al.* (2000) found that colectomy with ileorectal anastomosis (IRA), with lifelong surveillance of the rectal stump is a good treatment but there is still a 15% risk of rectal cancer.

Arvantis *et al.* (1990) found that the main cause of death for people with FAP is colorectal cancer, 58.2% (colon 44.5% and rectal 13.6%), followed by desmoid tumors, 10.9%, periampullary cancer, 8.2%, brain tumors, 7.3%, perioperative mortalities, 4.5%, adrenal carcinoma, 0.9%, and carcinomatosis, 0.9%. After prophylactic surgery (total proctocolectomy of total abdominal colectomy and ileorectal anastomosis) the causes of death change significantly. The main cause of death is now desmoid tumours, 30.6%, followed by periampullary cancer, 22.2%, perioperative mortalities, 8.3%, rectal cancer, 8.3%, adrenal cancer, 2.8% and carcinomatosis, 2.8%. It has been shown that there are no deaths caused by colon cancer after surgery and that, while there is still a risk of rectal cancer, there are fewer deaths caused by rectal cancer.

#### 4.3.6 Survival

Heiskanen *et al.* (2000) estimated that for all of those with FAP, 91% of propositi deaths are attributable to FAP, compared with 70% of call-up group deaths. They found that the call-up group had a significant survival advantage after surgery (colectomy or proctocolectomy). The mortality of the call-up group was found to be very similar to a comparable group in the general population. The majority of the deaths were due to colorectal cancer, 79%, with the second most common cause of death in FAP patients being rectal stump cancer.

Järvinen (1992) has shown that the life expectancy of those diagnosed through family screening is significantly better from the age of 31 years. Bülow (1986) agrees, estimating the cumulative survival rates ten years after diagnosis to be 42% for those diagnosed through symptoms and 97% for those diagnosed by way of family screening, while he estimates the survival rates as 94% for call-up patients

and 44% for probands in 2003.

## 4.4 Application to Insurance

Each of the quantities discussed above are of interest to the medical profession as they can be used to provide sufferers with information about their disorder and also to carry out further research.

These quantities are also of interest to the insurance industry and could have an effect on the premiums that would be charged for different types of insurance;

The onset rate and the survival rate will affect the premium in different ways depending on the age and the term of the policy sought. For example if the onset of a disease is most likely to occur between ages 20 and 30, the CI premium, as a percentage of the standard rate, would be higher at age 20 and 30 than at age 50, as at age 50 they are past the age at which they are most at risk.

There are gender effects to be taken into account when determining standard rate insurance premiums. If there is a gender effect on either the age at onset, severity of the disease or the survival then this will result in the premium, relative to the standard, being different for each sex.

# Chapter 5

## Tools and Methodology — Estimating Onset Rates

### 5.1 Onset Rate and Penetrance

As discussed, briefly, in Section 2.3.1 the intensity, or onset rate, is defined as the rate at which onset of a disease is suffered and is defined as a function of age  $x$  and denoted  $\mu(x)$ .

The penetrance is defined as:

$$F(x) = P(\text{Onset has occurred by age } x) \quad (5.1)$$

and is related to the intensity by:

$$F(x) = 1 - \exp\left(-\int_0^x \mu(t)dt\right). \quad (5.2)$$

When considering the effect of genetic testing on premium rates the onset rate is an important item of information. The onset rate is often determined through likelihood methods (discussed in Section 5.3).

### 5.2 Classical Survival Analysis

Survival analysis is based on the observation of a sample of  $N$  subjects, where each subject's lifetime is a random variable. We denote the lifetime from birth of the  $i$ th

subject  $X^i$ , and we assume that  $X^1, \dots, X^N$  are iid, each with distribution  $F_X(x)$  and density  $f_X(x)$ .

The hazard function can then be written as:

$$\mu(x) = \lim_{h \rightarrow 0^+} \frac{P[X^i \leq x + h | X^i > h]}{h} \quad (5.3)$$

that is, the instantaneous probability of death per unit time having survived to age  $x$ . It is the same as the intensity, or onset rate. The intensity is also given by:

$$\mu(x) = \frac{dF_X(x)/dx}{1 - F_X(x)} = \frac{f_X(x)}{1 - F_X(x)} = \frac{f_X(x)}{S_X(x)} \quad (5.4)$$

where  $S_X(x)$  is the survival function, defined as  $S_X(x) = 1 - F_X(x)$ . Hence:

$$S_X(x) = \exp\left(-\int_0^x \mu(t)dt\right) \quad (5.5)$$

or  $S(x) = \exp(-A(x))$  where  $A(x) = \int_0^x \mu(t)dt$  is defined to be the cumulative hazard.

We can use these results to compile a likelihood by first defining the random variables:

$$\tilde{X}^i = \text{highest age observed for the } i\text{th person}, \quad (5.6)$$

$$Z^i = \begin{cases} 1, & \text{Person } i \text{ dies at age } \tilde{X}^i \\ 0, & \text{Observation of person } i \text{ was censored at time } \tilde{X}^i. \end{cases} \quad (5.7)$$

Then the likelihood of the data contributed by the  $i$ th person can be written as:

$$L^i = S_X(\tilde{X}^i) \mu(\tilde{X}^i)^{Z^i} = \exp\left(-\int_0^{\tilde{X}^i} \mu(t)dt\right) \mu(\tilde{X}^i)^{Z^i}. \quad (5.8)$$

See Chapter II of Andersen *et al.* (1993) for more details.

When the event of interest is the onset of a disorder and not death, as in the case of critical illness insurance, lifetime can be defined as time free from the disorder.

Figure 5.1 illustrates the two-state model for critical illness insurance. A person can be in two possible states in this model; state 0 if they are healthy or state 1 if they have suffered onset of disease, this is the most basic model of onset. Here  $\mu_{01}(x)$  represents the transition intensity from state 0 to state 1. In this instance we estimate the probabilities that a person is in state 0 or state 1, given whatever information is known.

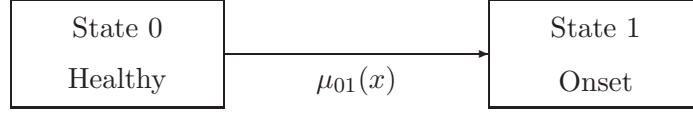


Figure 5.1: The two-state model for critical illness insurance.

### 5.3 Pedigree Likelihoods

Suppose we have a pedigree with  $N$  members, that each suffered onset or censoring at age  $x_i$ , and that the  $i$ th member has phenotype  $z_i = 1$ , affected, or  $z_i = 0$ , unaffected and has genotype  $u_i$ . Suppose also that there are  $M$  alleles at a single locus. Then the possible genotypes of the whole pedigree are the  $N$ -dimensional vectors of the set  $U = \{1, \dots, M\}^N$ .

Also for pedigree genotype  $v \in U$  let  $P[v] = P[u_i = v_i : i = 1, \dots, N]$ , where  $P[v]$  is determined through the allele frequencies and the mode of transmission, and let  $g_{v_i}(z_i)$  be the probability that the  $i$ th person, with genotype  $v_i$ , has phenotype  $z_i$ . The genotype probabilities,  $P[v]$ , are determined by applying the genetic model to the genotype probabilities of the founders. The genotype probabilities of the founders are the population frequencies of the  $M$  alleles.

Supposing that age of onset is variable, we take it to have density function  $\varphi_u(x)$  and survival function  $\Phi_u(x)$ , (where survival is defined here as absence of onset) conditional on genotype  $u$ . Also let  $x_i$  be the age at which the  $i$ th person's genotype  $z_i$  is observed.

Then Elston (1973) writes the likelihood of a pedigree as:

$$L = \sum_{v \in U} \left( \prod_i (z_i g_{v_i}(1) \varphi_{v_i}(x_i) + (1 - z_i) ((1 - g_{v_i}(1)) \Phi_{v_i}(x_i))) \right) P[v]. \quad (5.9)$$

where  $g_{v_i}(0) = 1 - g_{v_i}(1)$ .

The pedigree likelihood is made up of contributions from each person within the pedigree. There are four possible contributions a person can make to the likelihood:

Table 5.1: Non-zero terms in the likelihood (5.9) for the pedigree in Figure 5.2. Genotypes are  $u_i = 1 =$  unaffected and  $u_i = 2 =$  affected.  $f$  is the frequency of the mutation within the population

Genotype	Contribution to the Likelihood
(2,2,2,1,2,2,1,1)	$f \ 0.5^2 \ \varphi_2(x_1)\varphi_2(x_2)\varphi_2(x_3)\Phi_1(x_4)\Phi_2(x_5)\Phi_2(x_6)\Phi_1(x_7)\Phi_1(x_8)$
(2,2,2,2,2,2,1,1)	$f \ 0.5^2 \ \varphi_2(x_1)\varphi_2(x_2)\varphi_2(x_3)\Phi_2(x_4)\Phi_2(x_5)\Phi_2(x_6)\Phi_1(x_7)\Phi_1(x_8)$
(2,2,2,1,2,2,1,2)	$f \ 0.5^2 \ \varphi_2(x_1)\varphi_2(x_2)\varphi_2(x_3)\Phi_1(x_4)\Phi_2(x_5)\Phi_2(x_6)\Phi_1(x_7)\Phi_2(x_8)$
(2,2,2,2,2,2,1,2)	$f \ 0.5^2 \ \varphi_2(x_1)\varphi_2(x_2)\varphi_2(x_3)\Phi_2(x_4)\Phi_2(x_5)\Phi_2(x_6)\Phi_1(x_7)\Phi_2(x_8)$

A mutation carrier who has suffered onset of the disorder -  $\varphi_2(x_i)$

A mutation carrier who has not suffered onset of the disorder -  $\Phi_2(x_i)$

A non-mutation carrier who has not suffered onset of the disorder -  $\Phi_1(x_i)$

A non-mutation carrier who has suffered onset of the disorder -  $\varphi_1(x_i)$

The latter is not possible in the case of Huntington's Disease but is possible for other disorders.

### 5.3.1 Example

Using the pedigree in Figure 5.2 as an example we evaluate the likelihood above, see Table 5.1. Here  $M = 2$  and  $U = \{1, 2\}$ <sup>8</sup>.

When a person tests negative or positive for the mutation their genotype is determined to be 1 or 2, respectively.

Persons I:1, II:1 and II:2 have all suffered onset of Huntington's disease, which implies that they are mutation carriers, so their contributions to the likelihood are  $\varphi_2(x_1)$ ,  $\varphi_2(x_2)$  and  $\varphi_2(x_3)$ , respectively.

Persons II:4 and III:1 have both tested positive for the mutation but have not yet suffered onset of the disease so their contributions are  $\Phi_2(x_5)$  and  $\Phi_2(x_6)$ , respectively.

Person III:2 has tested negative for the mutation and has not suffered onset giving

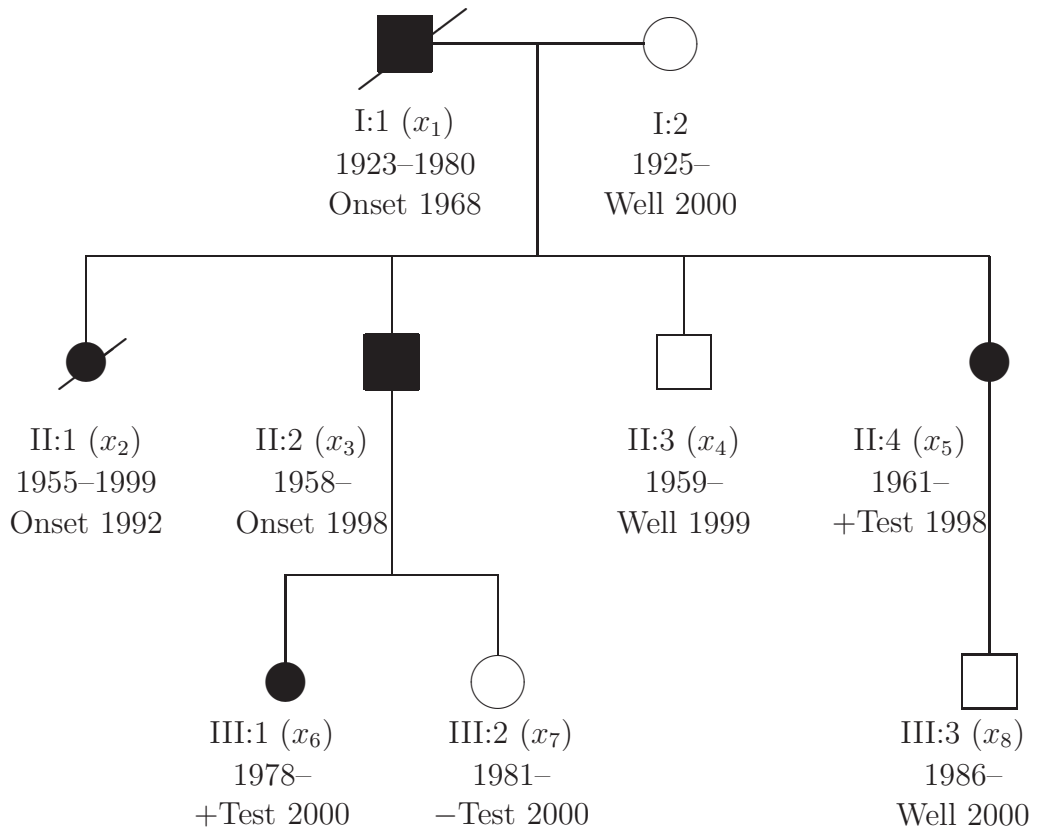


Figure 5.2: Three generations of a sample pedigree. Squares are males, circles are females, and a slash denotes death. Affected individuals are shown as filled squares/circles. The year of birth, death and other relevant information for each person is shown. Individuals are labelled  $X:Y$ , where ‘ $X$ ’ (Roman) labels the generation and ‘ $Y$ ’ (Arabic) the individual.  $x_i$  represents the age of person  $i$  at onset or censoring.



a contribution of  $\Phi_1(x_7)$ .

There are usually some people in the pedigree with unknown genotype. They are dealt with by assuming every combination of possible pedigree genotypes. Taking cases of onset and genetic tests into consideration, assuming here that onset implies positive genotype, only members II:3 (age  $x_4$ ) and III:3 (age  $x_8$ ) have unknown genotype. The four possible genotype combinations for these two members are:

Person II:3 negative and III:3 negative

Person II:3 positive and III:3 negative

Person II:3 negative and III:3 positive

Person II:3 positive and III:3 positive

These combinations result in the four possible likelihood contributions in Table 5.1. Summing over these possible combinations of the genotype gives the pedigree likelihood as:

$$L = 0.5^2 \varphi_2(x_1) \varphi_2(x_2) \varphi_2(x_3) \Phi_2(x_5) \Phi_2(x_6) \Phi_1(x_7) \left( \sum_{g=1}^2 \Phi_g(x_4) \sum_{g=1}^2 \Phi_g(x_8) \right). \quad (5.10)$$

## 5.4 Genetic Test Results

For those who are known to have a mutation the penetrance is taken as;

$$F(x) = 1 - \exp \left( - \int_0^x \mu_{01}(t) dt \right) \quad (5.11)$$

where  $\mu_{01}(x)$  is the onset rate, or transition intensity, introduced in Section 3.5 and shown in Figure 5.1 so

$$\mu_{01}(x) = \frac{F'(x)}{1 - F(x)}. \quad (5.12)$$

This gives the intensity for a person known to be a mutation carrier.

The survival function for a known mutation carrier (genotype  $u = 2$ ) is:

$$\begin{aligned} \Phi_2(x) &= \exp \left( - \int_0^x \frac{F'(t)}{1 - F(t)} dt \right) = \exp \left( - \int_0^x \frac{d}{dt} \log(1 - F(t)) dt \right) \\ &= \exp(\log(1 - F(x)) - \log(1 - F(0))) = 1 - F(x) \end{aligned} \quad (5.13)$$

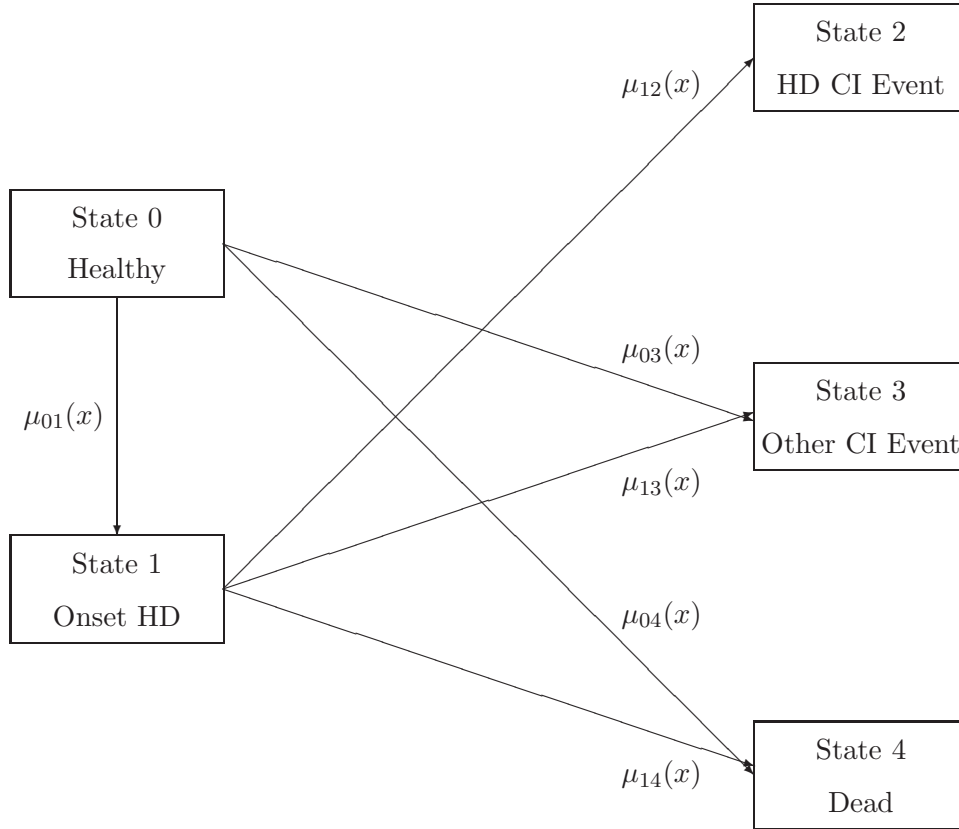


Figure 5.3: A multiple state model for HD in Critical Illness insurance.

and similarly the density of age at onset is:

$$\begin{aligned}
 \varphi_2(x) &= \exp\left(-\int_0^x \frac{F'(t)}{1-F(t)} dt\right) \left(\frac{F'(x)}{1-F(x)}\right) \\
 &= (1-F(x)) \left(\frac{F'(x)}{1-F(x)}\right) = F'(x).
 \end{aligned} \tag{5.14}$$

## 5.5 Huntington's Disease

### 5.5.1 Mutation Carrier

The Normal distribution is a good fit to the distribution of the age at onset of Huntington's Disease, as demonstrated by Newcombe (1981) and Smith (1998), so we will use the Normal distribution for the penetrance here;

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) dt \tag{5.15}$$

Which gives;

$$\varphi_2(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (5.16)$$

if onset occurs or;

$$\Phi_2(x) = \left(1 - \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) dt\right) \quad (5.17)$$

if it does not.

### 5.5.2 Non-Mutation Carriers

As non-mutation carriers (genotype  $u = 1$ ) are not at risk of Huntington's Disease they do not contribute to the onset rate.

Here;

$$\Phi_1(x) = 1 \text{ and } \varphi_1(x) = 0 \quad (5.18)$$

## 5.6 Breast and Ovarian Cancer

The likelihood is slightly more complex for breast and ovarian cancer. Here we have to determine the onset rate for both breast and ovarian cancer. In order for us to estimate the intensities we have to maximise the likelihood to estimate the parameters of the distributions of the age at onset of both BC and OC simultaneously. We use the idea of competing risks with independent lifetimes. We consider the time to development of breast cancer and the time to development of ovarian cancer to be bivariate random variables. We can then assume that the two random variables are independent conditional on genotype, it is possible that some factors may influence both, but this is not known. Therefore we can treat the probability of developing breast or ovarian cancer as a product of the individual probabilities of developing the disorders. While estimating the distribution of the age at onset for one disorder we treat the onset of the other disorder as a censoring event.

The likelihood also has to be altered to allow for the fact that neither ovarian or breast cancer are 100% penetrant. The penetrance values that will be used are taken from published papers, rather than estimated as additional parameters.

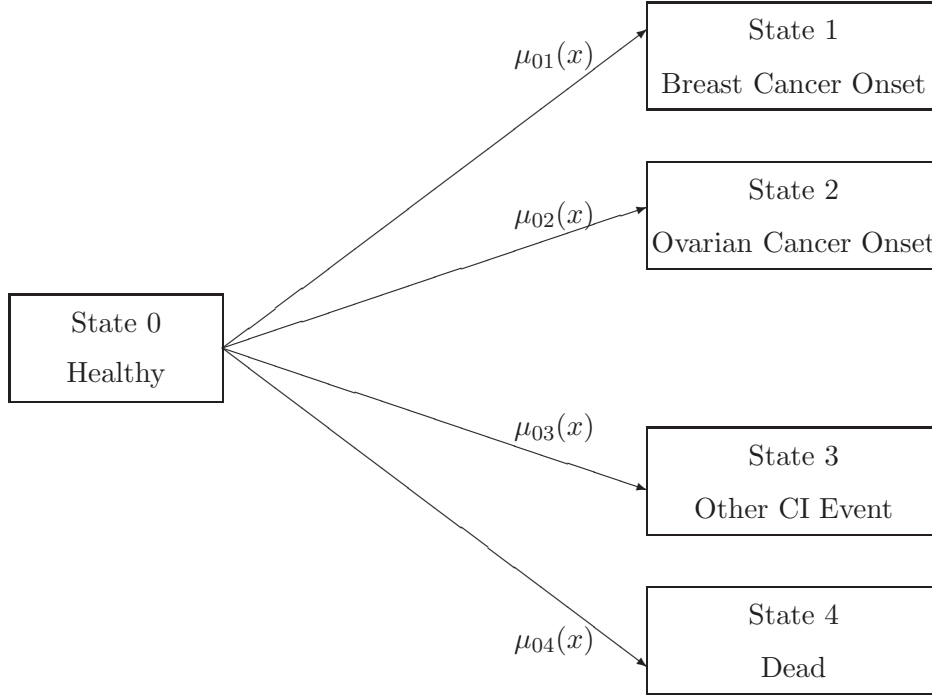


Figure 5.4: A multiple state model for BRCA1/BRCA2-related breast and ovarian cancer in Critical Illness insurance.

Here we only determine the onset rate for BRCA1 carriers as the data is not sufficient to enable us to estimate the onset rate for BRCA2 carriers.

### 5.6.1 Mutation Carriers

The distributions of age at onset of BRCA1/2 related breast and ovarian cancer are skewed as the excess risk falls at later ages. Therefore we will use the Gamma distribution for the age at onset distribution of both breast and ovarian cancer:

$$F^j(x) = \rho^j \frac{\beta^j \alpha^j}{\Gamma(\alpha^j)} \int_0^x t^{\alpha^j-1} e^{-\beta^j t} dt, \quad (5.19)$$

where  $j$  denotes the disorder,  $j = BC$  for breast cancer and  $j = OC$  for ovarian cancer and  $\rho^j$  is the penetrance.

This gives;

$$\varphi_2^{BC}(x) = \rho^{BC} \frac{\beta^{BC} \alpha^{BC}}{\Gamma(\alpha^{BC})} x^{\alpha^{BC}-1} e^{-\beta^{BC} x} \left( 1 - \rho^{OC} \frac{\beta^{OC} \alpha^{OC}}{\Gamma(\alpha^{OC})} \int_0^x t^{\alpha^{OC}-1} e^{-\beta^{OC} t} dt \right),$$

(5.20)

if onset of breast cancer occurs, given mutation carrier status, and

$$\varphi_2^{OC}(x) = \rho^{OC} \frac{\beta^{OC} \alpha^{OC}}{\Gamma(\alpha^{OC})} x^{\alpha^{OC}-1} e^{-\beta^{OC} x} \left( 1 - \rho^{BC} \frac{\beta^{BC} \alpha^{BC}}{\Gamma(\alpha^{BC})} \int_0^x t^{\alpha^{BC}-1} e^{-\beta^{BC} t} dt \right),$$

(5.21)

if onset of ovarian cancer occurs, given mutation carrier status.

If onset does not occur we have;

$$\begin{aligned}\Phi_2^{BC}(x) = \Phi_2^{OC}(x) = & \left(1 - \rho^{BC} \frac{\beta^{BC} \alpha^{BC}}{\Gamma(\alpha^{BC})} \int_0^x t^{\alpha^{BC}-1} e^{-\beta^{BC}t} dt\right) \\ & \times \left(1 - \rho^{OC} \frac{\beta^{OC} \alpha^{OC}}{\Gamma(\alpha^{OC})} \int_0^x t^{\alpha^{OC}-1} e^{-\beta^{OC}t} dt\right)\end{aligned}\quad (5.22)$$

The parameters for both breast and ovarian cancer will be maximised simultaneously using the methods outlined in Appendix B.

## 5.6.2 Non-Mutation Carriers

People who are known or assumed not to carry mutations are still at risk of ovarian and breast cancer so we need onset rates in respect of them. We take the following rates from Macdonald, Waters & Wekwete (2003a).

$$\begin{aligned}\mu_{01}(x) &= \frac{1}{\Gamma(8.7305)} 0.0742^{8.7305} e^{-0.0742x} x^{7.7305} & 0 \leq x < 53 \\ \mu_{01}(x) &= 0.00012 + 0.00018(x - 35) - 0.000005(x - 35)^2 & x \geq 53 \\ &+ 0.0000000529(x - 35)^3\end{aligned}\quad (5.23)$$

$$\begin{aligned}\mu_{02}(x) &= \frac{1}{\Gamma(6.92)} 0.035^{6.92} e^{-0.035x} x^{5.92} & 0 \leq x < 45 \\ \mu_{02}(x) &= 0.0001554 + 0.000029(x - 45) - 0.00000048(x - 45)^2 & x \geq 55\end{aligned}\quad (5.24)$$

with linear interpolation between ages 45 and 55.

The contribution to the likelihood is then:

$$\varphi_1^{BC}(x) = \exp\left(-\int_0^x \mu_{01}(t) dt\right) \mu_{01}(x) \exp\left(-\int_0^x \mu_{02}(t) dt\right), \quad (5.25)$$

if onset of breast cancer has occurred, and;

$$\varphi_1^{OC}(x) = \exp\left(-\int_0^x \mu_{02}(t) dt\right) \mu_{02}(x) \exp\left(-\int_0^x \mu_{01}(t) dt\right), \quad (5.26)$$

if onset of OC occurs.

If onset is not suffered the likelihood contribution would be:

$$\Phi_1^{BC}(x) = \Phi_1^{OC}(x) = \exp\left(-\int_0^x \mu_{01}(t) + \mu_{02}(t) dt\right). \quad (5.27)$$

## 5.7 Variance of Parameter Estimates

The variances and covariances of the MLEs are found by taking derivatives of the likelihood function. We will begin by outlining the method used for a twice-continuously differentiable one-dimensional function,  $f(x)$ .

Using Taylor's expansion, where  $f_x$  is the first derivative,  $f_{xx}$  the second and so on:

$$f(x+h) = f(x) + hf_x + \frac{1}{2}h^2f_{xx} + \frac{1}{6}h^3f_{xxx} + \dots \quad (5.28)$$

and

$$f(x-h) = f(x) - hf_x + \frac{1}{2}h^2f_{xx} - \frac{1}{6}h^3f_{xxx} + \dots \quad (5.29)$$

Then

$$f(x+h) - f(x-h) \simeq 2hf_x + \frac{1}{3}h^3f_{xxx}. \quad (5.30)$$

Therefore as  $h$  tends to zero, following the definition of the derivative:

$$f_x \simeq \frac{f(x+h) - f(x-h)}{2h}. \quad (5.31)$$

Applying the same method to  $f_x$  we get:

$$f_{xx} \simeq \frac{f(x-h) + f(x+h) - 2f(x)}{h^2}. \quad (5.32)$$

For a two-dimensional function, Taylor's expansion gives the following approximate derivatives:

$$f_x \simeq \frac{f(x+h_x, y) - f(x-h_x, y)}{2h_x}, \quad (5.33)$$

$$f_y \simeq \frac{f(x, y+h_y) - f(x, y-h_y)}{2h_y}, \quad (5.34)$$

$$f_{xy} \simeq \frac{[f(x+h_x, y+h_y) + f(x-h_x, y-h_y)] - [f(x+h_x, y-h_y) + f(x-h_x, y+h_y)]}{4h_xh_y}, \quad (5.35)$$

$$f_{xx} \simeq \frac{f(x+h_x, y) + f(x-h_x, y) - 2f(x, y)}{h_x^2}, \quad (5.36)$$

$$f_{yy} \simeq \frac{f(x, y+h_y) + f(x, y-h_y) - 2f(x, y)}{h_y^2}. \quad (5.37)$$

Using these derivatives the variance matrix can be estimated as

$$V = \begin{bmatrix} V(x) & Cov(x, y) \\ Cov(x, y) & V(y) \end{bmatrix} = - \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}^{-1} \quad (5.38)$$

from which we can construct confidence intervals or simulate values in the parameter space consistent with the sampling distribution.

## 5.8 Bootstrapping

Maximising the pedigree likelihood using the simplex method (Appendix B) provides us with estimates of the parameters of the distribution of age at onset, and of the variance matrix using the method outlined in Section 5.7.

The estimated onset rates will then be used in the calculation of premium rates. As the premium rate is an estimate ultimately based on the data it would be useful to know how reliable it is by looking at its sampling variance. Here we will use a bootstrapping method. The procedure is as follows:

- Estimate the onset rates, using a parametric model:

$$\text{Rate of Onset at age } x = \mu(x; \theta) \quad (5.39)$$

where  $\theta$  is the (vector) parameter and  $\hat{\theta}$  is its estimate.

- Estimate the sampling variance matrix  $Var(\hat{\theta})$ .
- Assume  $\hat{\theta}$  to have a multivariate normal sampling distribution and simulate  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(S)}$  as independently generated variates from the multivariate normal distribution with mean  $\hat{\theta}$  and variance  $Var(\hat{\theta})$ .
- Calculate the premium rate,  $g(\theta)^{(i)}$ , for each simulated  $\theta^{(i)}$ . Then the empirical distribution of  $g(\theta^{(i)})$  gives us the bootstrap estimate of the sampling distribution of the premium rate.

Information about the variance of onset rate estimates is not routinely available in published epidemiological studies and it is uncommon for data to be available



for actuarial investigations of genetic disorders. As a result the variance of the subsequent premium rates is often ignored. The availability of the data for this investigation enables the reliability of the premium rate estimates to be evaluated, in a robust statistical sense, which is one original contribution of this study.

## 5.9 Ascertainment Bias

Ascertainment bias arises through the fact that if a family is not identified as having the disease by geneticists or researchers then there is no chance of them being in the sample to be studied. The affected person, or persons, who first brought the pedigree to the attention of geneticists or researchers is known as the proband, or probands.

If the researcher chooses, or is aware of, only families with a higher number of cases then the onset rates estimated using these cases will be too high.

Many methods have been suggested in the published literature concerning the best way to deal with the problem of ascertainment bias. The general idea is that an adjustment should be made to the likelihood to account for the fact that without the proband, or probands being affected, it would be unlikely that the pedigree would come to the attention of the researchers.

Sometimes a pedigree is identified through a number of people from the same family being affected by the same disorder, called multiple ascertainment. This is often the case for breast and ovarian cancer, where one affected family member alone would probably not draw attention to the pedigree but a number of cases of the disorder, especially cases of early-onset BC and OC, would do so.

In the case of disorders such as Huntington's Disease, however, where genetic mutation is the only cause, the pedigree of an affected family is often identified through just one affected member, this is known as single ascertainment.

Thompson (1993) discusses adjusting for ascertainment by conditioning the likelihood on the probability that the pedigree was ascertained. In the case of single ascertainment, where there is only one proband, the probability that a pedigree is ascertained is proportional to the number of affected individuals in the pedigree

and the conditioning is equivalent to removing the proband from the likelihood.

Cannings and Thompson (1977) discuss ascertainment bias in sequential sampling. They showed that ascertainment bias can be avoided if the decision of who, in a pedigree, is to be examined next depends only on what has already been observed and that all of the people who are examined are included in the analysis, regardless of phenotype. This means that the decision to stop sampling from a particular pedigree should be based solely on the people who have already been examined.

Hodge and Vieland (1996) defined single ascertainment as any ascertainment scheme under which all pedigrees have the same probability of being ascertained independent of the pedigree size or structure. Under this definition they showed that the likelihood should be the probability of the data conditioned on the observed pedigree divided by a function  $p(\theta)$ , whether the sampling is proband-dependent (PD) or proband-independent (PI), where  $p(\theta)$  is a function of the genetic parameters but not the pedigree structure. They also showed that when  $p(\theta)$  is the prevalence of the trait being studied then the definition is equivalent to classical single ascertainment, as defined by Thompson (1993).

Ewens and Shute (1986) illustrated adjusting the likelihood by conditioning only on the part of the data which is relevant to ascertainment. They note that it may be difficult to decide what part of the data that is, but no assumption is made about how the probability of ascertainment is affected by data relevant to ascertainment.

The extension of the methods of Cannings and Thompson (1977) to multiple ascertainment were discussed by Boehnke and Greenberg (1984). They have shown that if ascertainment is complete then conditioning on the probands results in all of the affected individuals being removed from the analysis, where complete ascertainment is defined as when every individual has equal probability (100%) of being ascertained. They showed that this method is not necessarily correct as the resulting estimates may be asymptotically biased.

Epidemiological studies have used a variety of methods to adjust for ascertainment bias. The methods of Cannings and Thompson (1977) were used by Antoniou *et al.* (2006) who conditioned the likelihood of all of the phenotypic and genotypic information given all of the disease phenotypes and genotypic information up to

the point at which the first carrier was identified.

Antoniou *et al.* (2003), Ford *et al.* (1998) and Anglican Breast Cancer Study Group (2000) maximised the likelihood of the pedigree given the phenotypic and genotypic information of the proband. Other studies have adjusted for ascertainment by removing the proband from the data set, Bear *et al.* (1992), Bear *et al.* (1984) and Churchill *et al.* (1984).

Newcombe (1981) tested three adjustment methods; (1) Not removing any probands; (2) Removing one proband per pedigree; and (3) Removing the proband only when the pedigree was singly ascertained. They found that the resulting estimates of onset rates under (2) and (3) were very close, while estimates under (1) were slightly lower.

In the case of the pedigree data available to us here no probands are identified, nor do we know the ascertainment scheme. As a result we cannot adjust the likelihood accurately using the aforementioned methods. We will therefore use a variation of the ‘leave one out’ method to adjust for ascertainment bias.

For Huntington’s Disease we assume a pedigree is generally identified through a single proband and as we are unsure of who may be the proband in each pedigree we ‘choose’ one at random, as follows. For each pedigree, we identify all possible probands as being those that suffered onset of HD before the first genetic test was taken. After identification of all possible probands we will randomly remove a ‘proband’ from each pedigree and determine the parameters of the distribution of the onset rate  $\mu(x; \theta)$ . Repeating this random exclusion of probands  $D$  times leaves us with a sample of  $\hat{\theta}$  and the corresponding variance matrix  $Var(\hat{\theta})$ . For each of these  $D$  parameter sets we then simulate  $E$  independently generated variates, assuming  $\hat{\theta}$  has a multivariate normal sampling distribution, as in Section 5.8.

We then have  $D * E$  estimates and we can, again, use these estimates to determine the premium rates  $g(\theta)$ , then the empirical distribution of  $g(\theta^{(i)})$  gives us a bootstrap estimate of the sampling distribution of the premium rate, now given the adjustment for ascertainment.

Comparing these values to those obtained through bootstrapping without the randomly deleted proband, we will have a measure of how the chosen method of dealing

with ascertainment bias affects the results. The variance of the premium rate will incorporate the uncertainty about who the probands are as well as the uncertainty about the age at onset.

In the case of breast and ovarian cancer pedigrees we first carry out the same ‘leave-one out’ method using randomly selected probands to adjust for single ascertainment. However, breast and ovarian cancer are not caused solely by the BRCA1 and BRCA2 genes, so the way in which they come to the attention of the researchers is often through multiple cases in one family. So to adjust for multiple ascertainment we also randomly remove two of the possible probands from each pedigree. This will result in estimated onset rates after approximately adjusting for single and multiple ascertainment and will enable us to compare the results obtained through both adjustment methods.

# Chapter 6

## Tools and Methodology — Estimating Premium Rates

We will first introduce the methods used to calculate the premium rates in respect of breast and ovarian cancer. Some alterations have to be made to the model to use it for HD, we will discuss these changes in Section 6.5.

There is usually a 28-day survival clause in a critical illness insurance policy so that the insurer pays out only if a person survives 28 days after onset.

To define the actuarial model used we must first introduce the continuous-time Markov model.

### 6.1 The Markov Model

The insurance model is defined in terms of states and the transitions between those states, where the states usually include Healthy, CI event and Death, as illustrated in Figure 6.1. The transition intensities and transition probabilities govern movement between states  $j$  and  $k$ . The transition probability,  $p_{jk}(x, t)$ , is defined as the probability that a person in state  $j$  at age  $x$  will be in state  $k$  at age  $x + t$ . The transition intensity between states  $j$  and  $k$  at age  $x + t$  is denoted  $\mu_{jk}(x + t)$ .

If the transition is between the healthy state and the onset state then the transition intensity is the onset rate defined in Section 2.3.1, with  $j = 0$  and  $k = 1$ .

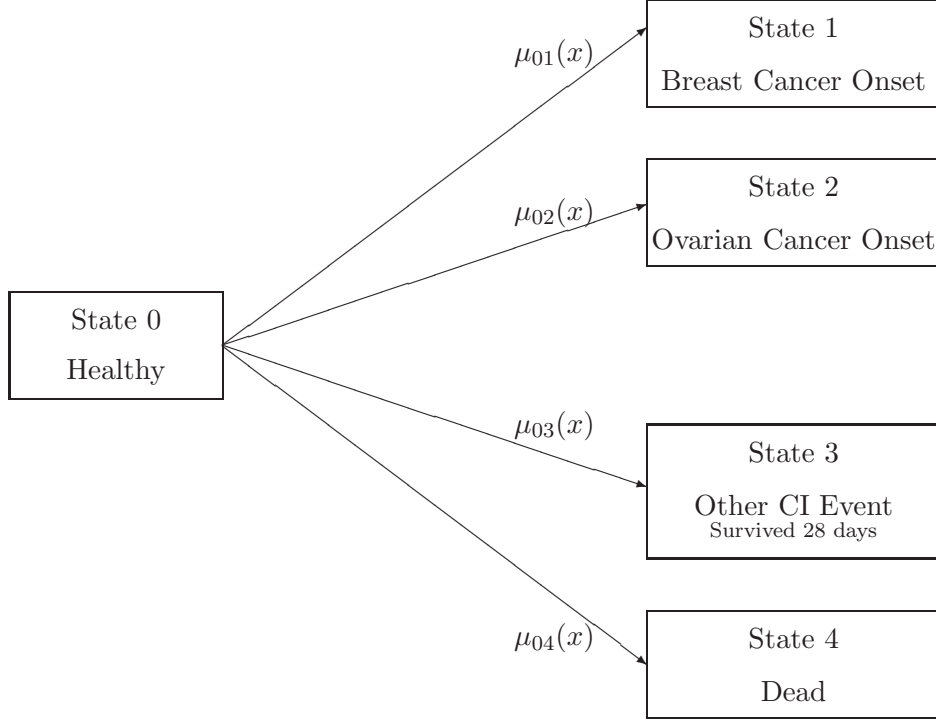


Figure 6.1: A multiple state model for BRCA1/BRCA2-related breast and ovarian cancer in Critical Illness insurance.

For BRCA1-related breast and ovarian cancer these transition intensities would be  $\mu_{01}(x)$  for breast cancer and  $\mu_{02}(x)$  for ovarian cancer (see Figure 6.1).

The transition probabilities and transition intensities are related by the Kolmogorov forward equations:

$$\frac{\partial}{\partial t} p_{jk}(x, t) = \sum_{l \neq k} p_{jl}(x, t) \mu_{lk}(x + t) - \sum_{l \neq k} p_{jk}(x, t) \mu_{kl}(x + t) \quad (6.1)$$

with boundary conditions  $p_{jk}(x) = 1_{j=k}$ .

We suppose that if a person is in state  $j$  at age  $x$ , then a premium is paid continuously at an annual rate of  $b_j(x)$  and on transition to state  $k$  a benefit of  $b_{jk}(x)$  will be paid, if  $k$  is a state that warrants a CI claim.

Denoting the force of interest as  $\delta$  the Expected Present Value (EPV) of the future discounted loss at age  $x + t$ ,  $V_j(x, t)$ , conditional on presence in state  $j$  is then related to the benefit and premium through Thiele's differential equations:

$$\frac{\partial}{\partial t} V_j(x, t) = \delta V_j(x, t) + b_j(x) - \sum_{k \neq j} \mu_{jk}(x, t) (b_{jk}(x, t) + V_k(x, t) - V_j(x, t)) \quad (6.2)$$

with boundary conditions  $V_{jk}(x, d) = 0$  at expiry of the policy, at some age  $x + d$ .

These equations can be solved numerically using methods found in any numerical analysis textbook (we use a fourth-order Runge-Kutta method with step-size 0.0005 year and a force of interest of  $\delta = 0.05$  per annum).

Solving the Kolmogorov forward equations gives us the transition probabilities, given all of the transition intensities. Then solving Thiele's differential equations gives the EPVs conditional on presence in any of the states.

## 6.2 The Model of Critical Illness Insurance Given Genetic Information

Figure 6.1 shows the Markov model for breast and ovarian cancer. Each relevant subgroup of the population will be represented by this model. As we concentrate only on the incidence of breast and ovarian cancer related to BRCA1 mutations there are three subgroups:

- (a) Persons who are not at risk.
- (b) Persons at risk, because of family history, but who do not have a BRCA1 mutation.
- (c) Persons at risk, because of family history, who do have a BRCA1 mutation.

The mutation frequencies give the proportions in the starting state of each subpopulation at birth. Table 6.1 shows these proportions.

## 6.3 Intensities

For subgroup (c) we determine the onset rate,  $\mu_{01}(x)$  and  $\mu_{02}(x)$  using the methods discussed in Chapter 5. The other transition intensities we use are those estimated by Gutiérrez & Macdonald (2004). These are as follows:

Table 6.1: Proportions of the whole population who are in each of the three subpopulations. Source: Gui *et al.* (2006).

Subpopulation	Proportion
(a)	Non BRCA1 carrier family 0.997671796
(b)	BRCA1 carrier family, applicant does not have mutation 0.001164102
(c)	BRCA1 carrier family, applicant has the mutation 0.001164102

### 6.3.1 Cancer

The intensities are based on cancer registrations between 1990 and 1992, from the Office of National Statistics (ONS, 1999).

For males:

$$\mu_{03c}(x) = \exp(-11.25 + 0.105x) \quad x < 51, \quad (6.3)$$

$$\mu_{03c}(x) = \exp(0.2591585 - 0.01247354x + 0.000191691x^2 - 8.952933 \times 10^{-7}x^3) \quad x \geq 60, \quad (6.4)$$

and between the ages of 51 and 60 linear interpolation is used.

For females:

$$\mu_{03c}(x) = \exp(-10.78 + 0.123x - 0.00033x^2) \quad x < 53, \quad (6.5)$$

$$\mu_{03c}(x) = -0.01545632 + 0.0003805097x \quad x \geq 53. \quad (6.6)$$

CI policies typically include a clause stating that after a CI event occurs the policy does not pay out until 28 days after diagnosis. The 28-day survival probability is not taken into account here as death is very unlikely within 28 days of diagnosis of cancer.

When considering BRCA1 mutation carriers we have to adjust the onset intensity as the onset rate of breast and ovarian cancer is already included in the model



through  $\mu_{01}(x)$  and  $\mu_{02}(x)$ , respectively. We do this by removing the population onset rate of breast and ovarian cancer given in Section 5.6.2:

$$\begin{aligned}\mu_{01}(x) &= \frac{1}{\Gamma(8.7305)} 0.0742^{8.7305} e^{-0.0742x} x^{7.7305} & 0 \leq x < 53 \\ \mu_{01}(x) &= 0.00012 + 0.00018(x - 35) - 0.000005(x - 35)^2 & x \geq 53 \\ &+ 0.0000000529(x - 35)^3\end{aligned}\quad (6.7)$$

$$\begin{aligned}\mu_{02}(x) &= \frac{1}{\Gamma(6.92)} 0.035^{6.92} e^{-0.035x} x^{5.92} & 0 \leq x < 45 \\ \mu_{02}(x) &= 0.0001554 + 0.000029(x - 45) - 0.00000048(x - 45)^2. & x \geq 55\end{aligned}\quad (6.8)$$

Then the onset intensity for cancer for BRCA1 mutation carriers,  $\mu_{03cm}(x)$ , is then:

$$\mu_{03cm}(x) = \mu_{03c}(x) - \mu_{01}(x) - \mu_{02}(x). \quad (6.9)$$

### 6.3.2 Heart Attack

The intensities were determined using numbers of first-ever cases of heart attacks between September 1991 and August 1992, from the Morbidity Statistics from the General Practice Survey (McCormick, Fleming and Charlton, 1995).

For males:

$$\mu_{03h}(x) = \exp(-13.2238 + 0.152568x) \quad x < 44, \quad (6.10)$$

$$\mu_{03h}(x) = -0.001245109 + 0.000315605x \quad x > 49, \quad (6.11)$$

with linear interpolation between ages 44 and 49

and for females:

$$\mu_{03h}(x) = \frac{0.598694}{\Gamma(15.6412)} \times 0.15317^{15.6412} \exp(-0.15317x) x^{14.6412}. \quad (6.12)$$

### 6.3.3 Stroke

The intensities were determined from Stewart *et al.* (1999).

For males:

$$\mu_{03s}(x) = \exp(-16.9524 + 0.294963x - 0.001904x^2 + 0.00000159449x^3), \quad (6.13)$$

and for females:

$$\mu_{03s}(x) = \exp(-11.1477 + 0.081076x). \quad (6.14)$$

### 6.3.4 28-day survival probabilities

The probability a person survives for 28 days following a heart attack given that they are aged  $x$  is denoted by  $s_{03h}(x)$ , similarly the probability of surviving 28 days following a stroke is denoted by  $s_{03s}(x)$ .

The corresponding 28-day mortality probabilities are then defined as;

$$q_{03h}(x) = 1 - s_{03h}(x) \quad (6.15)$$

and

$$q_{03s}(x) = 1 - s_{03s}(x). \quad (6.16)$$

The values of the mortality rates that we use are taken from Dinani *et al.* (2000);

$$s_{03s}(x) = (0.9 - 0.002x)/0.9 \quad (6.17)$$

for both males and females, while

$$q_{03h}(x) = 0.21 \quad (6.18)$$

at ages 20-80 for females only. Table 6.2 shows  $q_{03h}(x)$  for males.

### 6.3.5 Other CI Illness Claims

It was assumed by Macdonald, Waters and Wekwete (2003b) and Dinani *et al.* (2000) that other minor causes of CI claims amount to 15% of claims from heart attack, stroke and cancer. Therefore, in total,

$$\mu_{03}(x) = 1.15(\mu_{03c}(x) + s_{03h}(x)\mu_{03h}(x) + s_{03s}(x)\mu_{03s}(x)) \quad (6.19)$$

Table 6.2: 28-day mortality rates after a heart attack for males. Source: Dinani *et al.* (2000).

Age	$q_{03h}(x)$	Age	$q_{03h}(x)$	Age	$q_{03h}(x)$	Age	$q_{03h}(x)$
20-39	0.15	58-59	0.21	47-52	0.18	65-74	0.24
40-42	0.16	60-61	0.22	53-56	0.19	75-79	0.25
43-46	0.17	62-64	0.23	57	0.20	80+	0.26

### 6.3.6 Mortality

The mortality intensity,  $\mu_{04}(x)$ , is determined from population mortality rates (English Life Tables No. 15 - O.P.C.S., 1991, 1993a & 1993b). These rates were adjusted to exclude deaths from causes leading to CI claims.

## 6.4 Calculating Extra Premiums when Genotype is Known

Using the intensities in Section 6.3 we calculate the level net CI premiums for a level £1 sum assured for any age and term of policy. The premiums are calculated by solving the equations in Section 6.1, using a fourth-order Runge-Kutta procedure with a step size of 0.0005 year and a force of interest of  $\delta = 0.05$  per annum. We then present the premium rate for mutation carriers as a percentage of the premium rate for non-mutation carriers.

## 6.5 The Insurance Model for Huntington's Disease

The CI model for HD is shown in Figure 6.2, which now has a state ‘onset of HD’, from which a person can suffer the HD CI event, another CI event or death. As discussed previously the onset of HD is not a critical illness event. The clinically

defined event of onset of HD is taken to be the time when the first symptoms appear. These first symptoms can be hard to recognise, in fact time of onset may often be ascertained retrospectively, and would not result in a CI insurance claim as they are usually very mild. There is no data available to model the time at which the symptoms reach a stage that they would trigger a claim. There is, however, data available on the time to clinically defined onset of HD and time to death from HD, after onset; a critical illness claim would occur between these two events.

Harper defined three stages of HD, as seen in Table 4.1, using his own observations and suggested that these three stages may be equal in duration. Here we assume that the CI claim occurs as disease progression moves into either stage 2 or 3 of Harper's progression, defined in Section 4.1. In Section 3.5.2 we discussed papers concerning HD and insurance. In this section we note that Gutiérrez & Macdonald (2004) used an accelerated lifetime model to adjust for the fact that the HD CI event does not occur until some time after onset and some time before death. The time to CI claim is modelled as an acceleration of the time to death after onset.

We take the random variable  $X$  to be the time in years from onset until death,  $F_X(x)$  to be the duration-dependent distribution of the lifetime after onset and  $\mu_{14}(x)$  to be the associated intensity.

We multiply the timescale by a constant,  $\eta \geq 1$ , to obtain a new random variable,  $Y$  with intensity  $\mu_{12}(x)$  such that:

$$F_Y(x) = P[Y \leq x] = P[X \leq \eta x] = F_X(\eta x). \quad (6.20)$$

The intensities associated with  $X$  and  $Y$  are then related through:

$$\mu_{12}(x) = \eta \mu_{14}(\eta x). \quad (6.21)$$

As the median of  $Y$  is  $1/\eta$  times the median of  $X$  it is clear that giving  $\eta$  a value of 1.5 or 3 will then correspond to the critical illness event occurring, and therefore claims being paid, after 2/3 or 1/3 of the survival time, respectively. These values may be taken to represent Stage 3 and Stage 2 of Harper's progression, respectively, and it is of interest to determine premiums at both of these stages as different insurers may use different definitions of the CI event for HD.

The other intensities in the model are defined as follows;

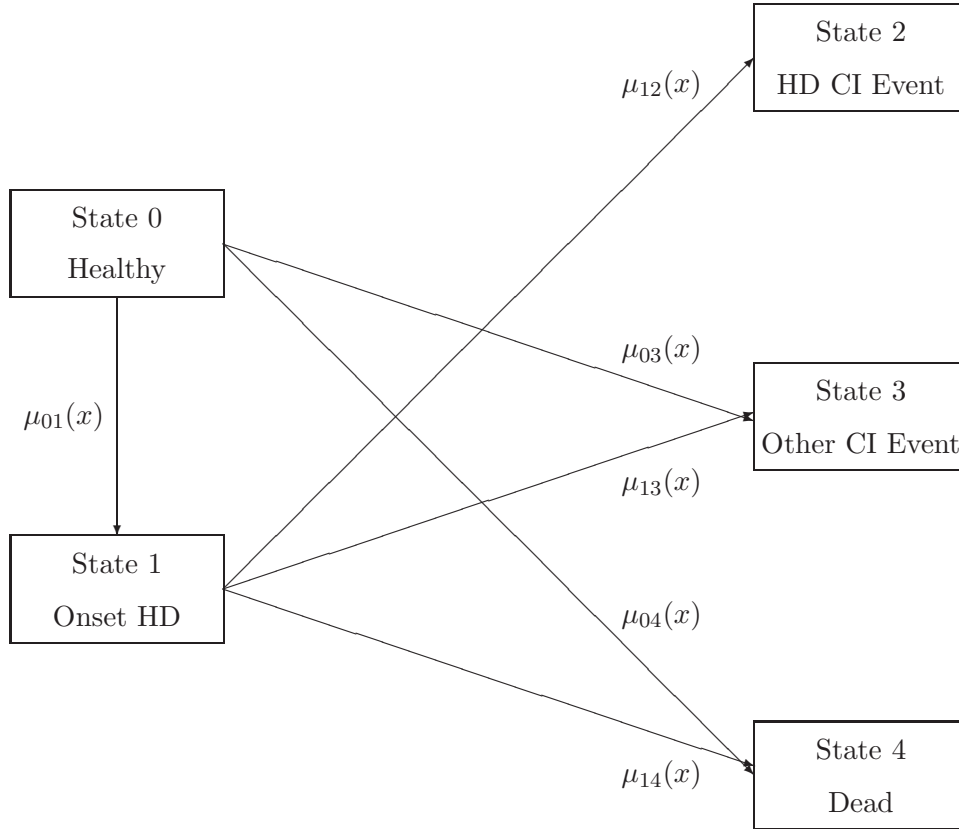


Figure 6.2: A multiple state model for HD in Critical Illness insurance.

We estimate the intensity  $\mu_{01}(x)$  using the data obtained from the pedigrees.

The estimates of the intensities  $\mu_{03}(x)$  and  $\mu_{04}(x)$  are the same as those for breast and ovarian cancer, without the adjustment for non-BRCA breast and ovarian cancer.

It is assumed here that the intensity  $\mu_{13}(x)$  is the same as for those who have not suffered onset of HD,  $\mu_{03}(x)$  as having HD would not significantly affect the likelihood of developing other critical illnesses.

For the intensity  $\mu_{14}(x)$  we use the same model of post-onset mortality as Gutiérrez & Macdonald (2002a) where the probability of surviving  $e$  years after onset of HD is  $S(e)$ .

For age at onset 20 - 34:

$$1 - S(e) = 0.174219^{4.11789} \int_0^e t^{3.11789} \exp^{-0.174219t} dt. \quad (6.22)$$

For age at onset 35 - 49:

$$1 - S(e) = 0.177225^{4.35046} \int_0^e t^{3.35046} \exp^{-0.177225t} dt. \quad (6.23)$$

For age at onset 50 or over:

$$1 - S(e) = 0.183372^{4.1465} \int_0^e t^{3.1465} \exp^{-0.183372t} dt. \quad (6.24)$$

For HD we also have to make a further alteration to the model which leads on from the accelerated lifetime model. Now that we are using duration-dependent survival rates after onset of HD the model becomes semi-Markov but certain calculations can be brought back into a Markov framework.

A reserve,  $V^{HD}(0, x + t)$ , has to be set up when a person moves from the insured healthy state to the onset state at age  $x + t$ . If the insurer pays out  $V^{HD}(0, x + t)$  as a sum assured at age  $x + t$ , equivalent to rewriting the risk at this point, then the policy value in all of the other states stays the same.

The final alteration is that the proportion of people from each sub-population is obviously different for HD, Table 6.3 shows these values.

Table 6.3: Proportions of the whole population who are in each of the three sub-populations. Source: Harper, Lim & Craufurd (2000)

Subpopulation	Proportion
Non HD carrier family	0.9996250
HD carrier family, applicant does not have mutation	0.0001875
HD carrier family, applicant has the mutation	0.0001875

# Chapter 7

## Huntington's Disease

### 7.1 Data

The data used for our study of HD were obtained from 60 pedigrees. The pedigrees were compiled by Jacki Needs and Professor Lindsay Prior from the records of the Institute of Medical Genetics at the University of Wales, Cardiff.

The pedigrees included from two to six generations. In most cases, the most recent generations contained persons too young to contribute any relevant information.

Apart from the blood relationships, we may have some or all of the following data, in respect of pedigree members:

- Calendar year of onset of HD.
- Calendar year in which a genetic test was carried out, and the result.
- Calendar year in which a person was last observed to be free of symptoms, called the age at censoring.
- Calendar year of death and cause of death.

We excluded those who suffered onset before the age of 20 or after the age of 60 for two reasons, one genetic and one actuarial. Juvenile HD (very early onset) is associated with very large numbers of CAG repeats. Including these cases may skew the estimated distribution of the age at onset. Their removal, and that of

very late-onset cases, is acceptable for actuarial purposes as the usual age range for insurance is 20 to 60.

In total we had data on 538 people from HD families. Eight of these people were suspected of having HD; they were not diagnosed clinically but their family suspected that they had suffered from HD. In respect of 74 people we had no information except year of birth and it was not known whether these people suffered onset of HD. These problems were dealt with in ways that will be discussed in Section 7.3.

## 7.2 An Example of a Pedigree

The best way to explain how the data were obtained from the pedigrees and to demonstrate how they were used is to give an example. Figure 7.1 is an example of a three-generation pedigree. This example is hypothetical and it is used to illustrate the features of the real pedigrees available. Here  $x_i$  represents the age at onset or censoring of the  $i$ th pedigree member.

**I:1** ( $x_1$ ) The grandfather, the founder of this pedigree, was diagnosed with HD when he was 40. This confirms that he has the huntingtin mutation.

**II:1** ( $x_2$ ) The most recent information available indicates that this male was well and free from HD at age 60, so the observation was censored at that age.

**II:2** ( $x_3$ ) This male was diagnosed with HD at age 44 and subsequently died at age 51.

**II:3** ( $x_4$ ) This female was diagnosed with HD at age 50.

**II:4** ( $x_5$ ) This male received a positive genetic test result at age 49, this was also their age at censoring.

**II:5** ( $x_6$ ) This woman died at age 26 (cause unknown), so the observation was censored at that age.

**III:1** ( $x_7$ ) This male tested negative for the huntingtin mutation and was censored at age 33.



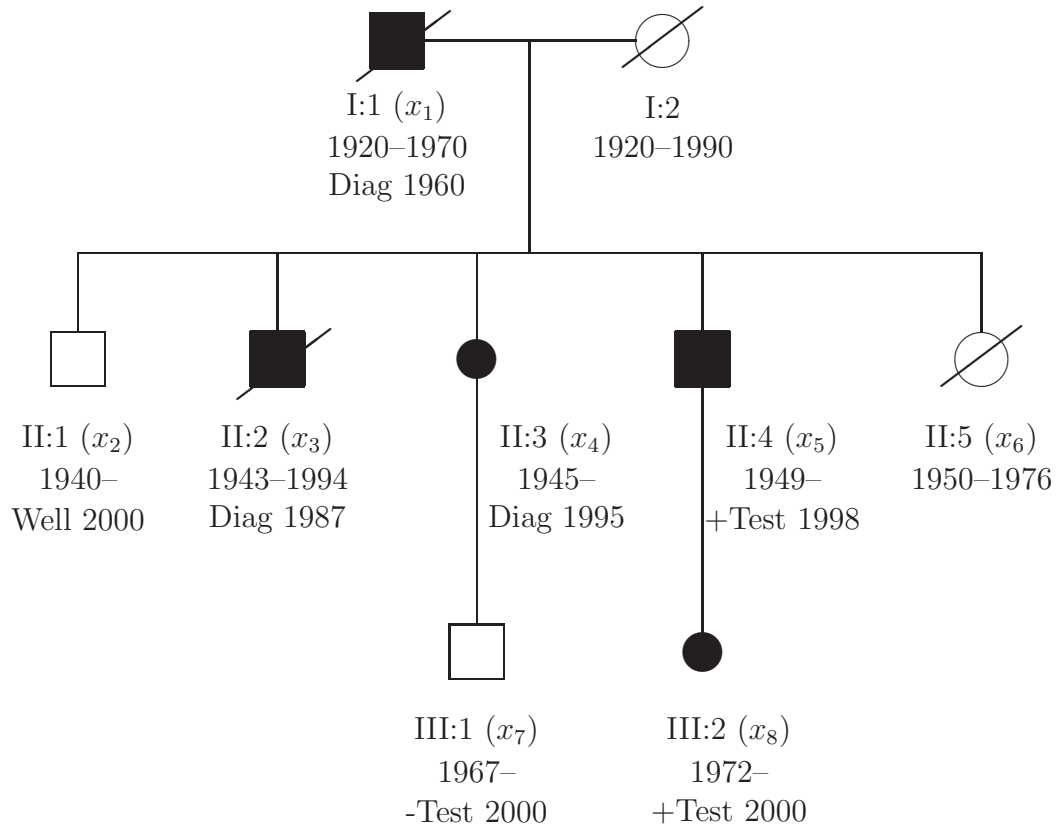


Figure 7.1: Three generations of a sample HD pedigree. Squares are males, circles are females, and a slash denotes death. Affected individuals are shown as filled squares/circles. The year of birth, death and other relevant information for each person is shown. Individuals are labelled  $X:Y$ , where ‘ $X$ ’ (Roman) labels the generation and ‘ $Y$ ’ (Arabic) the individual and  $x_i$  represents the age at onset or censoring.

**III:2 ( $x_8$ )** This woman tested positive for the huntingtin mutation and was censored at age 28.

## 7.3 Missing Data Problems

The main problem is that the dates of one or more of the key events may be missing in respect of one or more pedigree members.

The pedigree in Figure 7.1 has no missing information. To highlight the possible items of missing data we will introduce the following four problems with the pedigree:

**Problem 1.** Assume that person II:2 was diagnosed with HD postmortem (age at death was 51).

**Problem 2.** Assume that person II:2 suffered onset in their 40's.

We therefore do not know at what age this person suffered onset of HD.

**Problem 3.** Assume that person II:1 was born in 1940 and no other information is available. We therefore have no information about this person except their year of birth.

**Problem 4.** Assume that the relatives suspected that person II:3 suffered onset of HD at age 50. Therefore this person has not been diagnosed medically but the family believe that they did suffer from HD.

Our approach is to deduce or postulate upper and lower limits for the missing dates, and to estimate onset rates under each assumption to give a range of results. In practice, the range of results so obtained is acceptably small.

(a) If evidence of onset was found only at a postmortem examination then the age at death is an upper bound of possible onset times, and we will take ten years before death to be a lower bound.

– Possible solution to Problem 1.

*The upper and lower limits for the age at onset for Person II:2 would be 51 and 41, respectively.*

(b) A person may only be known to have suffered onset between certain ages. These ages are their own lower and upper bounds for the age of onset.

– Possible solution to Problem 2.

*Upper and lower bounds for the age at onset for Person II:2 would be 40 and 50, respectively.*

(c) If we know only the year of birth there are three ways to proceed.

(i) Assume that the person was alive and free from HD in the last calendar year in which their pedigree was observed (Assumption A1).

- Possible solution to Problem 3.  
*Assume Person II:1 alive and free from HD in the year 2000, age 60.*
- (ii) Exclude them from the sample (Assumption A2).
  - Possible solution to Problem 3.  
*Exclude only Person II:1 from the analysis.*
- (iii) Remove the entire sibship of which they are a member (Assumption A3).
  - Possible solution to Problem 3.  
*Exclude the whole pedigree from the analysis.*
- (d) In some cases HD had not been diagnosed by a medical practitioner but relatives suspected that a pedigree member had developed HD. In this case three assumptions are made:
  - (i) Assume that the pedigree member did develop HD at the time stated by the family (Assumption B1).
    - Possible solution to Problem 4.  
*Assume Person II:3 suffered onset of HD at age 50.*
  - (ii) Assume that the pedigree member did not develop HD and observation is censored at the time stated by the family (Assumption B2).
    - Possible solution to Problem 4.  
*Assume Person II:3 was alive and free from HD at age 50.*
  - (iii) Exclude the pedigree member from the sample (Assumption B3).
    - Possible solution to Problem 4.  
*Exclude Person II:3 from the analysis.*

Table 7.1 shows how each assumption should affect the estimated rate of onset. Making assumptions A1 and B1 together will have little or no effect on the estimated age at onset. Making assumption A1 will result in an overestimate while B1 will underestimate the age at onset, which implies that making both assumptions should result in an estimate very close to the one which would be achieved if there were no missing data.

Table 7.1: The assumptions made about missing or indeterminate data and how each assumption should affect the rate of onset. A1–A3 and B1–B3 are the labels referred to in Section 7.3.

Assumptions	Effect on Age at Onset
A1 & B1	Not significant
A2 & B1	Underestimate
A3 & B1	Underestimate
A1 & B2	Overestimate
A1 & B3	Overestimate

## 7.4 Results — Onset Rates

Table 7.2 shows the results under all the assumptions given in Table 7.1.

The first row of the table illustrates the maximum likelihood estimates of the mean  $m$  and standard deviation  $\sigma$  of the Normal distribution of age at onset, using all available data, excluding only those who suffered onset before age 20 and after age 60 and making Assumptions A1 and B1. The remainder of the table shows the results obtained under the assumptions defined in Section 7.3.

The table is presented in two parts, based on assumptions of ‘earliest possible age at onset’ and assumptions of ‘latest possible age at onset’ when dealing with missing data. It is clear that the assumptions have the expected effect on the estimated age at onset distribution.

The estimates for mean age at onset obtained here, when assuming ‘earliest age at onset’ lie in the interval (42.64, 43.98) and fall in the range of estimates from the studies in Table 4.2. The estimates obtained assuming ‘latest possible age at onset’ lie in the interval (43.71, 45.42) and are, for the most part, slightly higher than those estimated in previous studies.

All of our estimated means are lower than the estimate of 45.96 years obtained by Newcombe (1981).

Our age at onset distributions, although slightly higher than most previous studies

Table 7.2: The Normal model and sensitivity analyses. The parameters  $\hat{m}$  and  $\hat{\sigma}$  are the mean and standard deviation of the age at onset, respectively, and  $N$  is the sample size. The sampling variance of each estimate is shown in parentheses. Assumptions A1–B3 are described in Section 7.3.

‘Earliest Possible Age At Onset’				
<b>Assumption</b>	$N$	$\hat{m}$	$\hat{\sigma}$	$Cov(\hat{m}, \hat{\sigma})$
A1 & B1	538	43.67 (0.982)	13.26 (0.509)	0.229
A2 & B1	538	43.09 (0.962)	13.21 (0.502)	0.197
A3 & B1	397	42.64 (1.113)	13.27 (0.581)	0.212
A1 & B2	538	43.98 (1.062)	13.43 (0.551)	0.264
A1 & B3	538	43.84 (1.052)	13.40 (0.546)	0.257

‘Latest Possible Age At Onset’				
<b>Assumption</b>	$N$	$\hat{m}$	$\hat{\sigma}$	$Cov(\hat{m}, \hat{\sigma})$
A1 & B1	538	45.04 (1.166)	14.52 (0.603)	0.275
A2 & B1	538	44.41 (1.149)	14.47 (0.596)	0.238
A3 & B1	397	43.71 (1.350)	14.64 (0.704)	0.263
A1 & B2	538	45.42 (1.275)	14.80 (0.660)	0.320
A1 & B3	538	45.27 (1.266)	14.76 (0.655)	0.312

have estimated, are plausible. For the calculation of premium rates we will use the mean and standard deviation estimated under assumptions A1 and B1, assuming both ‘earliest’ and ‘latest’ ages at onset, that is; 43.67 and 13.26, and 45.04 and 14.52, respectively (all in years).

Figure 7.2 shows the distribution of age at onset assuming earliest and latest age at onset. Also shown in the figure are the distributions obtained assuming a Normal distribution for the mean and standard deviations found by Reed *et al.* (1958) and Wendt *et al.* (1959). These two studies represent the highest and lowest mean age at onset in Table 4.2, 35.3 and 43.97, respectively. The shape of the distributions of age at onset obtained through maximising the pedigree likelihood (assuming both

‘earliest’ and ‘latest’ age at onset) is very similar to that of both Bell (1934) and Wendt *et al.* (1959). There are slight differences in the position of the distributions, resulting from differences in the estimates of the mean parameter  $\hat{m}$ .



Figure 7.2: Distribution of age at onset of Huntington’s Disease estimated from the pedigree data, assuming both ‘earliest’ and ‘latest age at onset’. The figure also shows the distributions of age at onset of HD obtained assuming a Normal distribution for the mean and standard deviation found by Bell (1934) and Wendt *et al.* (1959).

## 7.5 Results – Premium Rates

As described in Section 5.8 we carried out bootstrapping using the estimates obtained in Section 7.4 to obtain the distribution of the premium rate, expressed as a percentage of the OR premium rate.

Simulated values of  $m$  and  $\sigma$  were obtained from the multivariate normal distribution,  $MN(a, b)$ , where  $a$  is the mean vector and  $b$  is the covariance matrix.

For ‘earliest possible age at onset’ we used:

$$a = (43.67, 13.26) \text{ and } b = \begin{pmatrix} 0.982 & 0.229 \\ 0.229 & 0.509 \end{pmatrix},$$

and for the ‘latest possible age at onset’

$$a = (45.04, 14.52) \text{ and } b = \begin{pmatrix} 1.166 & 0.275 \\ 0.275 & 0.603 \end{pmatrix},$$

Given the 1,000 bootstrapped estimates of the premium rate we then obtained estimates of the standard deviation of the sampling distribution of the premium rate estimate using the methods outlined in Section 5.8.

Estimates of these quantities for ‘earliest age at onset’, for a combination of ages and terms, are shown in Table 7.3 and Table 7.4, with the CI event defined as progression to Harper’s Stage 2 and Stage 3, respectively. Table 7.5 and Table 7.6 show the same results for ‘latest age at onset’.

The bootstrap-estimated sampling distributions of the premium rates obtained for an insurance policy term of 10 years are illustrated in Figures 7.3 and 7.4 (assuming ‘earliest’ and ‘latest’ ages at onset, respectively). The figures show the premium rate sampling distributions of both males and females, with the CI event defined as either Stage 2 or Stage 3 of Harper’s progression.

Figures 7.5 and 7.6 also show the sampling distributions of the premium rates but for a variety of policy terms, assuming the age at which the person purchased the insurance policy was 20 years. The four plots shown in each figure have differing scales as the actual ranges of the sampling distributions of the premium rate differ so much, depending on the definition of the CI event and the sex of the applicant.

We make the following comments:

- (a) The definition of the CI event affects the insurance premium significantly. Figures 7.3 and 7.4 show that, although the sampling distributions of the premium rate are very similar in shape for both Stages 2 and 3, there is a big difference in the average premium rate. When Stage 3 of Harper’s progression is defined as the CI event, the premium is significantly reduced.

- (b) Figures 7.5 and 7.6 show that the definition of the CI event and the sex of the applicant have a significant effect on the sampling distribution of the premium rate, particularly for shorter policy terms.
- (c) Given that a premium of about 300 – 350% of the standard rate will result in an applicant being declined insurance it is clear that insurance would only be offered to those at age 50 for a term of 10 years, when the CI event is defined as Harper’s Stage 3.
- (d) Gutiérrez and Macdonald (2003) estimated the premium, as a percentage of the standard rate, for different CAG repeat lengths (36 – 50). Their estimates for those with 50 CAG repeats are slightly lower than ours but they follow the same pattern. They found, however, that insurance could be offered for a number of combinations of CAG repeat lengths, terms and ages. Comparing their results for different CAG repeat lengths with our results highlights the fact that those with a lower number of CAG repeats would pay lower insurance premiums than if they were underwritten on the basis of family history alone.
- (e) The estimates of the standard deviation of the bootstrapped premium rate estimates give us an idea of the sampling variability of the estimated insurance premium. It is clear, from both the tables and the plots, that there is more variance in the premium rate for the lower ages and terms. This variance decreases as the term of the policy and the age when insurance is purchased increases.
- (f) Whether we take the ‘earliest’ or ‘latest’ possible age at onset does not make a big difference to the resulting sampling distributions of the premium rate, seen by comparing Figures 7.3 and 7.4, and Figures 7.5 and 7.6.



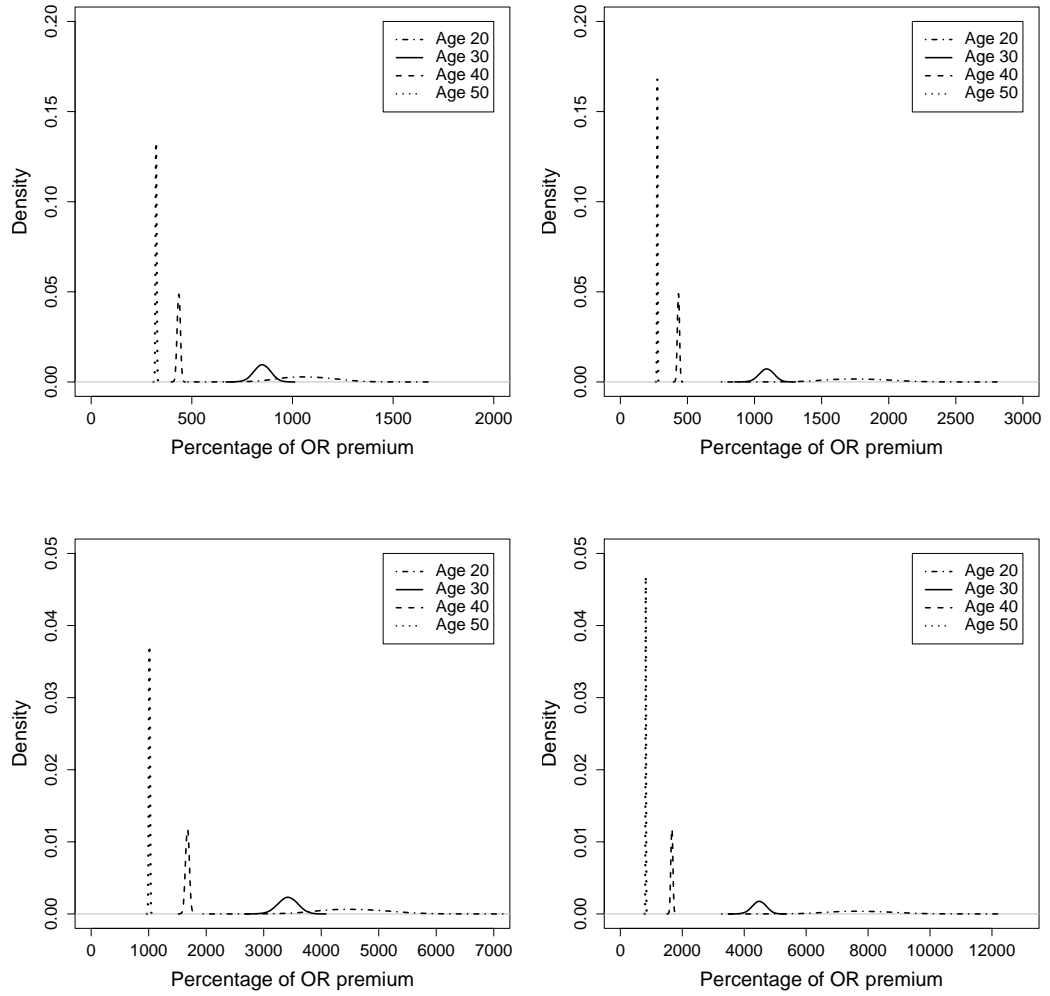


Figure 7.3: Sampling distribution of the premium rate for a range of ages, assuming ‘earliest possible age at onset’ and an insurance term of 10 years. The scales of the  $x$  and  $y$  axes are not the same for all four plots.

Top Left: Female with CI event defined as Harper’s Stage 3.

Top Right: Male with CI event defined as Harper’s Stage 3.

Bottom Left: Female with CI event defined as Harper’s Stage 2.

Bottom Right: Male with CI event defined as Harper’s Stage 2.

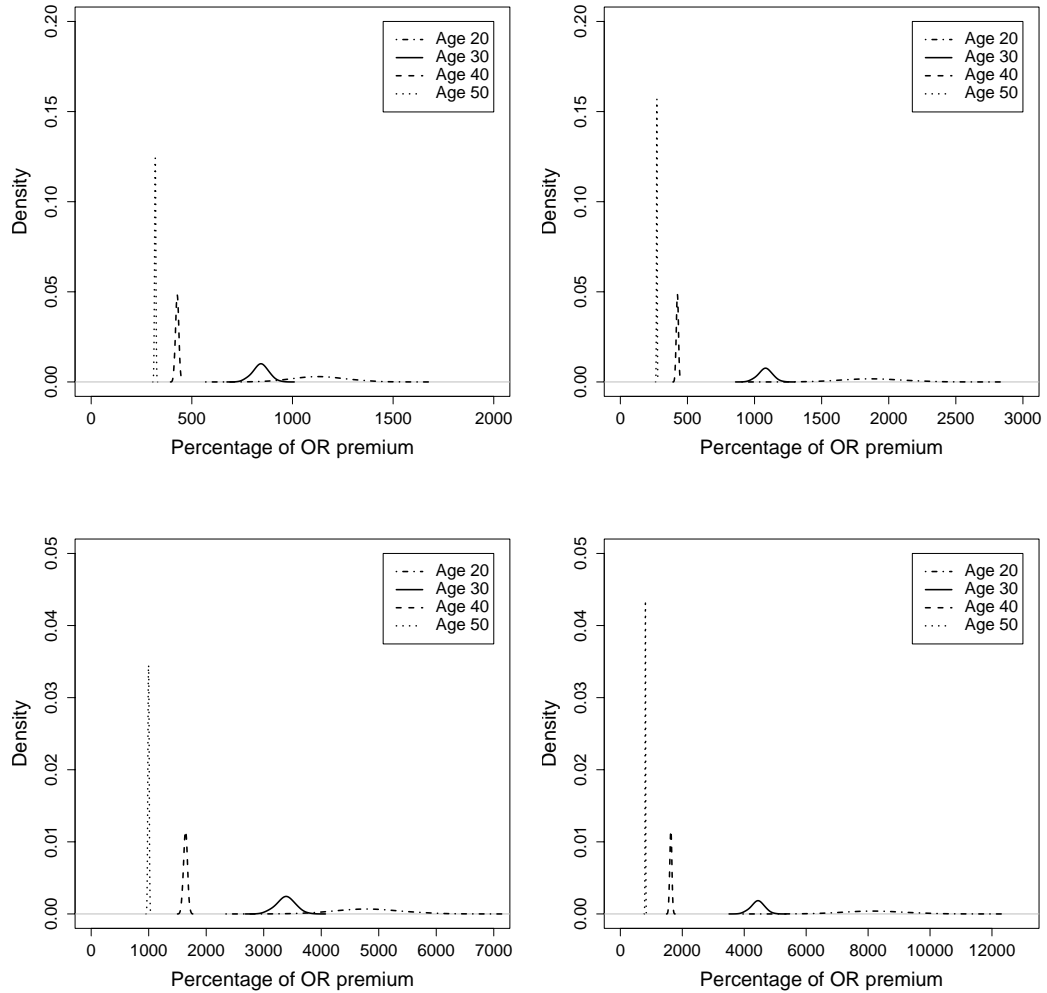


Figure 7.4: Sampling distribution of the premium rate for a range of ages, assuming ‘latest possible age at onset’ and an insurance term of 10 years. The scales of the  $x$  and  $y$  axes are not the same for all four plots.

Top Left: Female with CI event defined as Harper’s Stage 3.

Top Right: Male with CI event defined as Harper’s Stage 3.

Bottom Left: Female with CI event defined as Harper’s Stage 2.

Bottom Right: Male with CI event defined as Harper’s Stage 2.

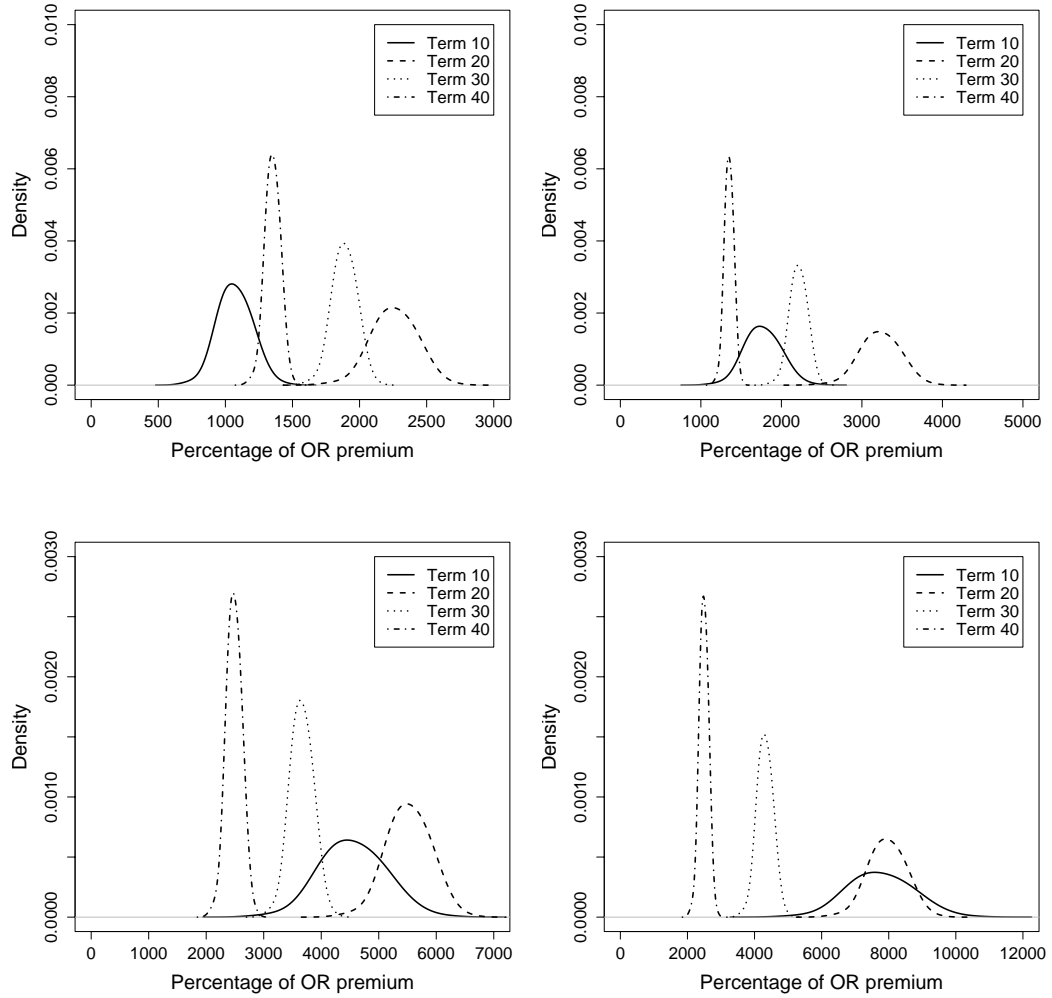


Figure 7.5: Sampling distribution of the premium rate for a range of policy terms, assuming ‘earliest possible age at onset’ and the age at which the policy is taken out is 20 years old. The scales of the  $x$  and  $y$  axes are not the same for all four plots.

Top Left: Female with CI event defined as Harper’s Stage 3.

Top Right: Male with CI event defined as Harper’s Stage 3.

Bottom Left: Female with CI event defined as Harper’s Stage 2.

Bottom Right: Male with CI event defined as Harper’s Stage 2.

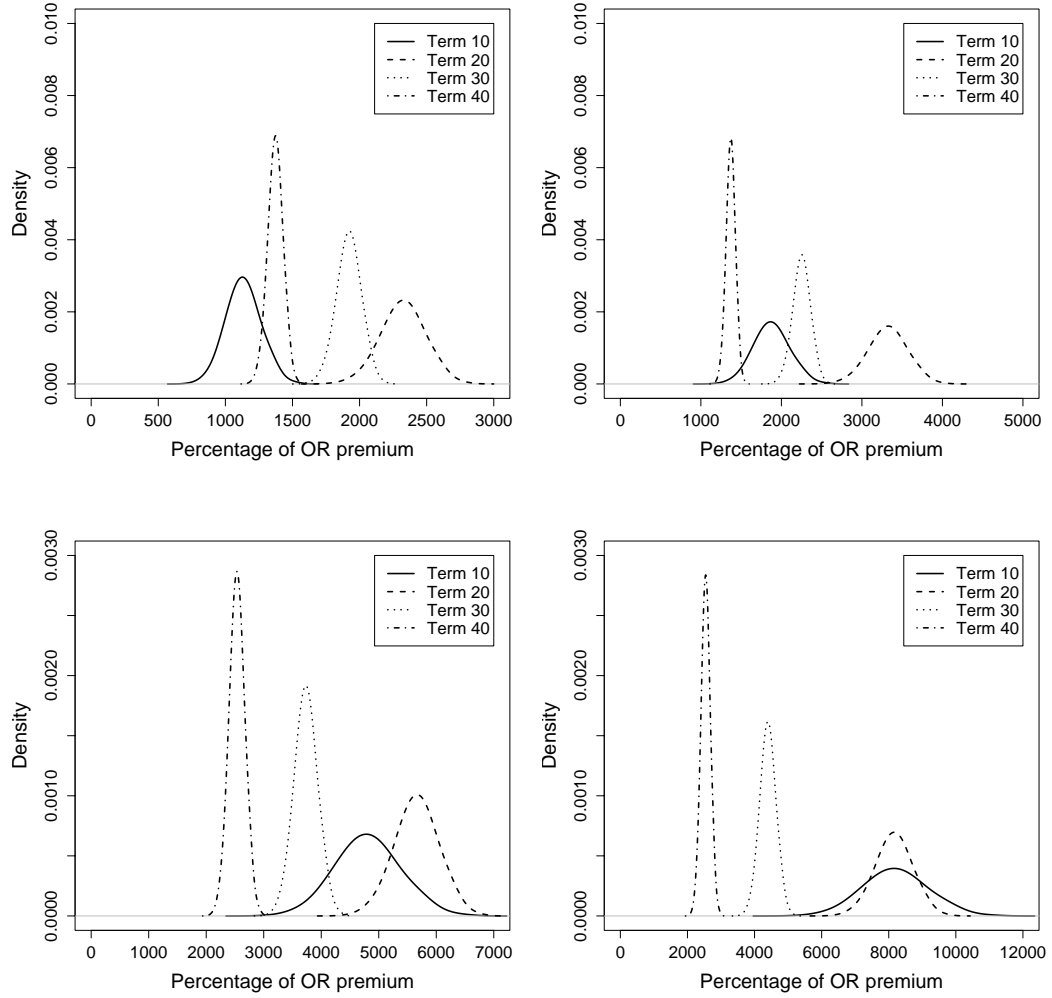


Figure 7.6: Sampling distribution of the premium rate for a range of policy terms, assuming 'latest possible age at onset' and the age at the which policy is taken out is 20 years old. The scales of the  $x$  and  $y$  axes are not the same for all four plots.

Top Left: Female with CI event defined as Harper's Stage 3.

Top Right: Male with CI event defined as Harper's Stage 3.

Bottom Left: Female with CI event defined as Harper's Stage 2.

Bottom Right: Male with CI event defined as Harper's Stage 2.

Table 7.3: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for the OR class, for claims at Stage 2 of Harper’s progression. Using parameters obtained given assumptions A1 & B1, and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	4,549	555.5
		20	5,522	377.3
		30	3,658	197.3
		40	2,485	131.9
	30	10	3,417	158.3
		20	3,077	81.0
		30	2,129	51.9
		10	1,673	30.4
	40	20	1,510	16.0
		10	1,013	9.6
Male	20	10	7,752	955.5
		20	7,956	546.9
		30	4,310	233.9
		40	2,493	133.2
	30	10	4,481	209.2
		20	3,374	89.2
		30	2,020	49.5
		10	1,663	30.2
	40	20	1,326	14.1
		10	821	7.6

Table 7.4: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for OR class, for claims at Stage 3 of Harper's progression. Using parameters obtained, given assumptions A1 & B1, and assuming 'earliest possible age at onset'. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	1,071	127.0
		20	2,260	165.3
		30	1,887	90.4
		40	1,352	55.6
	30	10	849	38.3
		20	1,298	33.0
		30	1,048	19.3
		40	435	7.3
	50	20	641	5.6
		10	323	2.7
Male	20	10	1,770	218.3
		20	3,228	239.5
		30	2,212	107.1
		40	1,350	56.0
	30	10	1,089	50.6
		20	1,416	36.3
		30	991	18.3
		40	433	7.3
	50	20	567	4.9
		10	276	2.1

Table 7.5: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 2 of Harper's progression. Using parameters obtained given, assumptions A1 & B1, and assuming 'latest possible age at onset'. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	4,813	547.8
		20	5,660	369.6
		30	3,732	194.5
		40	2,535	130.0
	30	10	3,382	156.9
		20	3,055	81.3
		30	2,116	51.9
		10	1,639	31.7
	50	20	1,492	16.8
		10	997	10.3
Male	20	10	8,205	942.2
		20	8,155	535.7
		30	4,398	230.6
		40	2,544	131.3
	30	10	4,435	207.3
		20	3,350	89.5
		30	2,007	49.5
		10	1,629	31.6
	50	20	1,311	14.8
		10	809	8.2

Table 7.6: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 3 of Harper's progression. Using parameters obtained given, assumptions A1 & B1, and assuming 'latest possible age at onset'. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	1,135	125.5
		20	2,325	160.4
		30	1,921	88.0
		40	1,373	54.2
	30	10	842	37.8
		20	1,290	33.1
		30	1,043	19.3
		40	427	7.6
	50	20	634	6.0
		10	318	2.9
Male	20	10	1,879	215.9
		20	3,323	232.4
		30	2,251	104.3
		40	1,371	54.7
	30	10	1,080	49.9
		20	1,407	36.4
		30	986	18.4
		40	425	7.5
	50	20	562	5.2
		10	272	2.3



### 7.5.1 Ascertainment Bias

The range of estimates of  $m$  obtained after adjusting for ascertainment bias, by randomly removing one proband from each pedigree are shown in Table 7.7. The range of the estimated  $m$  values is quite small and they are very similar to the estimates used for obtaining the premium rate without adjustment for ascertainment bias, (Section 7.4). There are some differences, however, in the covariance matrices associated with the estimates, in particular the variance of  $m$  ranges from 0.8 to 1.3 and the variance of  $\sigma$  ranges from 0.4 to 0.7, when assuming ‘latest age at onset’ and 0.8 to 1.0 and 0.4 to 0.5, respectively, when assuming ‘earliest age at onset’.

We then simulate from each of these estimated normal distributions, as explained in Section 5.9. The range of the estimates of  $m$  and  $\sigma$ , being used for estimation of the premium rates, is wider than when no ascertainment adjustment is carried out, see Table 7.8.

Table 7.7: The minimum and maximum bootstrapped values of  $m$  when adjusting for ascertainment bias, where  $m$  is the mean of the age at onset.

‘Earliest Possible Age At Onset’		‘Latest Possible Age At Onset’	
Min $m$	Max $m$	Min $m$	Max $m$
42.7	44.4	45.0	46.8

Table 7.8: The minimum and maximum of the bootstrapped values of  $m$  carrying out the random sampling from the multivariate normal distributions, where  $m$  is the mean of the age at onset.

‘Earliest Possible Age At Onset’		‘Latest Possible Age At Onset’	
Min $m$	Max $m$	Min $m$	Max $m$
39.1	47.6	41.2	50.6

Tables 7.9 – 7.12 show the premium rates, as a percentage of the standard rate, when the onset rate is adjusted for ascertainment bias.

The estimated premiums when an ascertainment adjustment is made, given as percentages of the OR premium, are slightly lower than when no adjustments are made. The only exceptions to this occur for those taking out the policy at age 50 for a term of 10 years, with the rate being slightly higher or equal after adjusting for ascertainment bias. The estimated mean age at onset, after adjustment for ascertainment bias, seems to be slightly higher than when no adjustment is made. For most of the ages and policy terms this would, and does, result in a lower premium rate. However, when no adjustment is made the risk begins to fall at a slightly earlier age than when adjustment is made and this results in the premium rate for age 50 being slightly higher when adjusted for ascertainment bias.

The standard deviation of the premium rate is higher when adjusted for ascertainment bias, this will be partly due to the fact that the variance in the estimates of the parameters of the distribution of age at onset is a lot higher after ascertainment adjustment.

When Harper's Stage 2 is defined as the CI event, the premium rate follows the same pattern as when no adjustment is made for ascertainment bias (Tables 7.3 and 7.5) .

When Stage 3 is used as the CI event the only insurable cases are still those at 50 years of age for a term of 10 years, but a number of cases have a premium closer to the 350% limit.

Figure 7.7 illustrates the difference that adjusting for ascertainment bias can make to the premium rate sampling distributions. These are shown for females only as, although the actual premium rate values for males are quite different, the relationship between the estimated sampling distributions with and without adjustment for ascertainment bias is very similar. For the same reason we show only the results obtained assuming 'latest possible age at onset'. The plots show that there are big differences in the the sampling distributions of the premium rates, with little overlap between the distributions with and without ascertainment adjustment. Although for HD this difference has no effect on the overall conclusions it is clear that it is important to adjust for ascertainment bias as it does have a large effect on the premium rate sampling distributions.

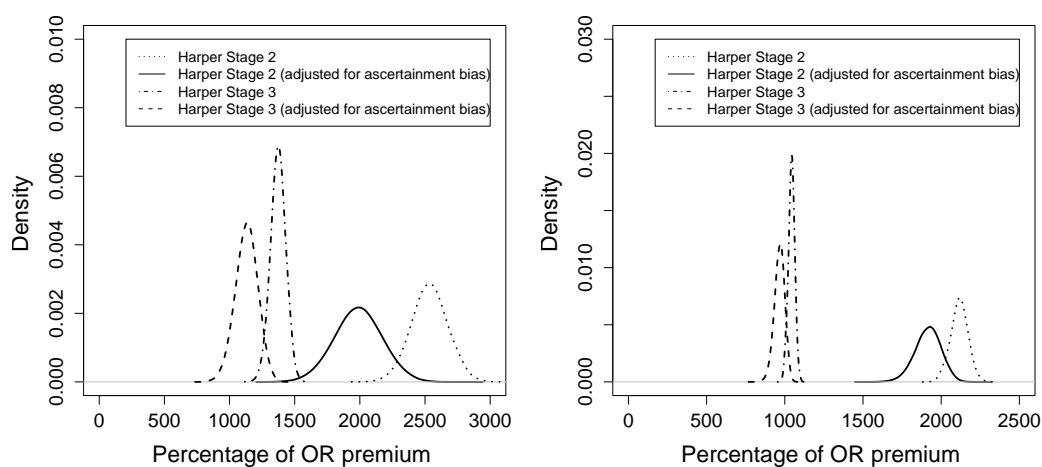


Figure 7.7: The sampling distributions for the premium rate assuming either Harper’s Stage 2 or Stage 3 as the CI event, with and without adjustment for ascertainment bias. The scales are not the same for both plots. Left: The distributions for Females of age 20 for a policy term of 40 years. Right: The distributions for Females of age 30 for a policy term of 30 years.

Table 7.9: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 2 of Harper's progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming 'earliest possible age at onset'.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	3,126	581.3
		20	4,517	457.2
		30	3,140	228.1
		40	2,144	151.1
	30	10	3,098	198.8
		20	2,916	104.1
		30	2,025	65.7
		10	1,662	35.8
	40	20	1,504	19.0
		10	1,023	10.8
Male	20	10	5,287	995.7
		20	6,513	660.4
		30	3,705	270.2
		40	2,145	153.1
	30	10	4,058	263.1
		20	3,199	113.9
		30	1,923	62.3
		10	1,652	35.8
	40	20	1,322	16.6
		10	830	8.6

Table 7.10: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 3 of Harper’s progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming ‘earliest possible age at onset’.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	742	130.2
		20	1,809	203.3
		30	1,642	114.3
		40	1,200	68.4
	30	10	769	47.9
		20	1,2231	43.7
		30	1,009	25.3
		40	432	8.7
	50	20	639	6.7
		10	325	3.1
Male	20	10	1,212	224.3
		20	2,583	293.0
		30	1,924	134.7
		40	1,199	69.1
	30	10	986	63.0
		20	1,344	47.5
		30	955	23.9
		40	430	8.6
	50	20	566	5.8
		10	278	2.4

Table 7.11: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 2 of Harper’s progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming ‘latest possible age at onset’.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,660	630.6
		20	4,048	550.6
		30	2,911	269.4
		40	1,988	179.1
	30	10	2,773	243.2
		20	2,736	133.5
		30	1,915	82.3
		10	1,577	43.7
	40	20	1,459	23.2
		10	999	12.6
Male	20	10	4,503	1,084.6
		20	5,842	773.4
		30	3,426	316.0
		40	1,994	179.2
	30	10	3,631	321.3
		20	3,002	142.6
		30	1,816	77.9
		10	1,568	43.4
	40	20	1,282	20.2
		10	810	10.0

Table 7.12: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks, for claims at Stage 3 of Harper’s progression. These estimates were obtained after adjusting for ascertainment bias, using parameters obtained given, assumptions A1 & B1, and assuming ‘latest possible age at onset’.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	646	140.7
		20	1,617	238.2
		30	1,522	139.6
		40	1,130	83.8
	30	10	695	57.7
		20	1,157	56.5
		30	965	32.9
		40	412	10.4
	50	20	623	8.3
		10	319	3.5
Male	20	10	1,038	241.2
		20	2,295	347.5
		30	1,779	165.5
		40	1,127	83.9
	30	10	885	76.1
		20	1,260	62.5
		30	913	31.3
		40	410	10.3
	50	20	552	7.3
		10	272	2.8

## 7.6 Conclusions

A number of factors will determine whether or not a person with a known HD mutation would be offered critical illness insurance.

Although insurance would not be offered for the majority of ages and policy terms the possibility of being able to take out insurance is greatly affected by the definition of the CI event used by the insurer. If Stage 2 of Harper's progression is the CI event used by an insurer then an applicant would not be offered insurance, regardless of age and the term of the policy they were seeking. If, however, Stage 3 of Harper's progression is defined as the CI event a person may be offered insurance for a policy term of 10 years when taking out insurance at 50 years old only.

Adjusting for ascertainment bias, by removing randomly selected 'probands' from each pedigree, has an effect on the sampling distribution of the premium rate. There are differences in the location and shape of the distribution with the distributions after adjusting for ascertainment bias having a slightly lower mean than when no adjustment was made and a higher standard deviation. This difference appears to be slightly greater when the CI event is defined as Stage 2 of Harper's progression rather than Stage 3. Although the effect of adjusting for ascertainment bias does not make any difference to the overall conclusions concerning insurance for those known to carry HD mutations there is a clear effect on the premium rate sampling distributions.

As described in Section 3.2, an application to GAIC for a specific genetic test and specific insurance product must include evidence of the actuarial significance of the genetic test. The results obtained here allow us to conclude that there is indeed a high probability that carriers of the HD mutation suffer onset of Huntington's Disease and that the critical illness event, either Stage 2 or 3 of Harper's progression, will occur at such an age that the premium rate that would be charged would result in the carrier being declined critical illness insurance at almost all ages and policy terms. However, as demonstrated by Gutiérrez and Macdonald (2003), using the extra information of CAG repeat lengths, not available for the pedigrees in this investigation, could result in insurance being offered at more ages and policy terms.



The main advantage of the work carried out here is that determining premium rate sampling distributions and adjusting for ascertainment bias, using actual pedigree data, results in the strengthening of any conclusions made. This will have significant relevance to GAIC, in terms of the evidence needed for any future applications to use a genetic test result to calculate insurance premiums, especially in the case of disorders which are less penetrant, and not as clear-cut as Huntington's Disease.

# Chapter 8

## Breast and Ovarian Cancer

### 8.1 Data

Thirty of the BRCA1/2 related breast and ovarian cancer pedigrees were obtained from the records of the Institute of Medical Genetics at the University of Wales, Cardiff. The other twenty-four were obtained from Southampton by our collaborators Jacki Needs and Professor Lindsay Prior.

Ten of the pedigrees were known to have the BRCA2 mutation while the rest were known to have BRCA1 mutations. As there was not enough data to estimate the onset rate of breast and ovarian cancer for BRCA2 carrying pedigrees we will consider only those with the BRCA1 mutation. This leaves us with 44 pedigrees.

For both breast and ovarian cancer we are using the Gamma distribution for the distribution of the age at onset and we will therefore have four parameters to estimate.

From the 44 BRCA1 pedigrees there were 270 people for whom sufficient data were available. Males were not included in the analysis unless they were known to have a BRCA1 mutation. As with HD there were a number of people for whom all information was missing with the exception of their year of birth. When all sibships containing any of these people were removed (Assumption A3 in Section 7.3) we were left with only 172 people.

The information obtained from these pedigrees was the same as for HD:

- Calendar year of onset.
- Calendar year in which a genetic test was carried out, and the result.
- Calendar year in which a person was last observed to be free of symptoms, called the age at censoring.
- Calendar year of death and cause of death.

## 8.2 Results – Onset Rates

As noted in Section 5.6 we will use previously published penetrance estimates for breast and ovarian cancer in the likelihoods to estimate the distribution of age at onset. So as to assess the effect of the assumed penetrance on the estimated age at onset distribution we will use results from three studies, see Table 8.1.

Table 8.1: The penetrance estimates that we use in the estimation of the onset rates of breast and ovarian cancer and the studies they were obtained from.

Reference	Breast Cancer	Ovarian Cancer
Lalloo <i>et al.</i> (2006)	84%	60%
Antoniou <i>et al.</i> (2006)	72%	38%
Antoniou <i>et al.</i> (2003)	65%	39%

Tables 8.2 – 8.4 show the results for all of the assumptions given in Section 7.3 for each set of penetrance estimates. Here we use assume that both ovarian and breast cancer follow a Gamma distribution with parameters  $\alpha$  and  $\beta$ . The first row of each table illustrates the results using all of the available data. The Gamma distribution obtained using the parameter estimates in the tables will be multiplied by the corresponding penetrance estimates to obtain the distribution of the age at onset. In this dataset the people for whom no further information is available are assumed to be alive and still at risk in the last calendar year in which the sibship was observed.

The estimates of  $\alpha$  and  $\beta$  obtained for breast cancer are consistently higher than for ovarian cancer regardless of the penetrance estimates used. There is no significant difference in the results obtained under either assumption A1 or A2 although carrying out the analysis under assumption A3 does increase the parameter estimates significantly.

For the estimation of the premium rates we will use the parameters, and variances, obtained through assumption A1 at both ‘earliest’ and ‘latest’ possible year. To obtain the actual distribution of the age at onset we multiply the gamma distribution by the corresponding penetrance estimates. As the results obtained using the penetrance estimates of both Antoniou *et al.* (2003) and Antoniou *et al.* (2006) are practically the same for ovarian cancer and quite close for breast cancer we will carry out the premium rate estimation using Antoniou *et al.* (2006) and Lalloo *et al.* (2006) only. Figures 8.1 and 8.2 show the estimated distributions of the age at onset for breast and ovarian cancer given the three different penetrance estimates, under assumption A1, and demonstrate how close the estimated distributions actually are.

Figure 8.3 shows the estimated distributions of age at onset for both breast and ovarian cancer, using the penetrance estimates of Lalloo *et al.* (2006) as well as the distributions obtained by Macdonald *et al.* (2003a) by fitting truncated Gamma distributions to the estimated incidence rates of Ford *et al.* (1998) for BRCA1 carriers. For both breast and ovarian cancer the distribution of age at onset estimated here is very similar in shape to that used by Macdonald *et al.* (2003a). There are, however, differences in the location. The study used by Macdonald *et al.* (2003a) found that breast and ovarian cancer were diagnosed slightly later, on average, (about 48 and 55, respectively) than we have in our data (44 and 46, respectively).

Table 8.2: The Gamma model and sensitivity analyses using the penetrance estimates of Lalloo *et al.* (2006).  $\alpha$  and  $\beta$  are the parameters of the Gamma distribution and  $N$  is the sample size. The variance of each estimate is shown in parentheses. Assumptions A1–A3 are described in Section 7.3.

‘Earliest Possible Age At Onset’						
Assumption	Disorder	$N$	$\alpha$	$\beta$	$Cov(\alpha, \beta)$	
A1	BC	270	16.36 (0.93)	0.41 (0.0007)	0.0224	
A2	BC	270	16.42 (0.38)	0.41 (0.0003)	0.0091	
A3	BC	172	19.28 (0.57)	0.48 (0.0005)	0.0137	
A1	OC	270	14.66 (1.11)	0.37 (0.0008)	0.0269	
A2	OC	270	14.69 (0.89)	0.37 (0.0006)	0.0215	
A3	OC	172	17.23 (1.09)	0.43 (0.0008)	0.0261	

‘Latest Possible Age At Onset’						
Assumption	Disorder	$N$	$\alpha$	$\beta$	$Cov(\alpha, \beta)$	
A1	BC	270	18.35 (0.43)	0.45 (0.0004)	0.0102	
A2	BC	270	18.40 (0.84)	0.45 (0.0006)	0.0120	
A3	BC	172	22.94 (0.44)	0.55 (0.0004)	0.0102	
A1	OC	270	16.42 (0.51)	0.40 (0.0004)	0.0122	
A2	OC	270	16.45 (0.67)	0.40 (0.0005)	0.0159	
A3	OC	172	20.49 (2.83)	0.50 (0.002)	0.0665	

Table 8.3: The Gamma model and sensitivity analyses using the penetrance estimates of Antoniou *et al.* (2006).  $\alpha$  and  $\beta$  are the parameters of the Gamma distribution and  $N$  is the sample size. The variance of each estimate is shown in parentheses. Assumptions A1–A3 are described in Section 7.3.

‘Earliest Possible Age At Onset’						
Assumption	Disorder	$N$	$\alpha$	$\beta$	$Cov(\alpha, \beta)$	
A1	BC	270	15.56 (0.33)	0.38 (0.0003)	0.0078	
A2	BC	270	15.59 (1.86)	0.38 (0.0012)	0.0442	
A3	BC	172	18.72 (1.05)	0.46 (0.0008)	0.0249	
A1	OC	270	14.71 (0.85)	0.36 (0.0006)	0.0202	
A2	OC	270	14.73 (0.85)	0.36 (0.0006)	0.0203	
A3	OC	172	17.55 (1.06)	0.43 (0.0004)	0.0252	

‘Latest Possible Age At Onset’						
Assumption	Disorder	$N$	$\alpha$	$\beta$	$Cov(\alpha, \beta)$	
A1	BC	270	17.34 (0.67)	0.41 (0.0005)	0.0158	
A2	BC	270	17.38 (0.97)	0.41 (0.0007)	0.0228	
A3	BC	172	21.73 (2.10)	0.52 (0.0013)	0.0489	
A1	OC	270	16.38 (1.31)	0.39 (0.0008)	0.0306	
A2	OC	270	16.40 (0.53)	0.39 (0.0004)	0.0123	
A3	OC	172	20.50 (1.50)	0.49 (0.0010)	0.0351	

Table 8.4: The Gamma model and sensitivity analyses using the penetrance estimates of Antoniou *et al.* (2003).  $\alpha$  and  $\beta$  are the parameters of the Gamma distribution and  $N$  is the sample size. The variance of each estimate is shown in parentheses. Assumptions A1–A3 are described in Section 7.3.

‘Earliest Possible Age At Onset’						
Assumption	Disorder	$N$	$\alpha$	$\beta$	$Cov(\alpha, \beta)$	
A1	BC	270	15.70 (5.86)	0.38 (0.0025)	0.1319	
A2	BC	270	15.75 (0.53)	0.38 (0.0004)	0.0125	
A3	BC	172	18.56 (0.54)	0.45 (0.0005)	0.0127	
A1	OC	270	14.77 (0.35)	0.36 (0.0002)	0.0053	
A2	OC	270	14.79 (2.17)	0.36 (0.0014)	0.0517	
A3	OC	172	17.51 (0.74)	0.43 (0.0006)	0.0176	

‘Latest Possible Age At Onset’						
Assumption	Disorder	$N$	$\alpha$	$\beta$	$Cov(\alpha, \beta)$	
A1	OC	270	17.52 (1.71)	0.42 (0.001)	0.0401	
A2	OC	270	17.56 (0.32)	0.42 (0.0003)	0.0076	
A3	OC	172	21.97 (0.85)	0.52 (0.0006)	0.0199	
A1	OC	270	16.47 (0.70)	0.39 (0.0005)	0.0165	
A2	OC	270	16.48 (0.84)	0.39 (0.0006)	0.0197	
A3	OC	172	20.59 (0.26)	0.49 (0.0003)	0.0061	

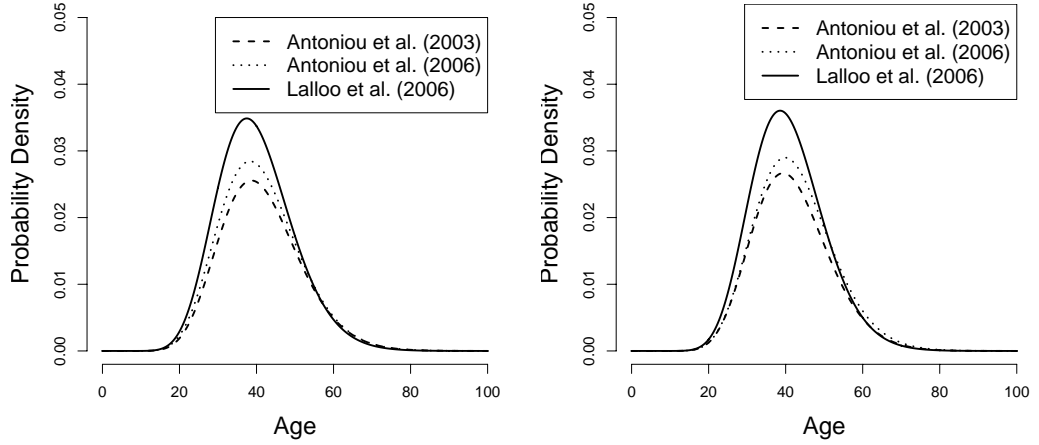


Figure 8.1: Distributions of age at onset for Breast Cancer obtained using the penetrance estimates of Antoniou *et al.* (2006), Antoniou *et al.* (2003) and Lalloo *et al.* (2006); Left: Assuming ‘Earliest age at onset’ and Right: Assuming ‘Latest age at onset’.

### 8.3 Results – Premium Rates

As for HD we carried out bootstrapping using the estimates obtained in Section 8.2 to obtain the distribution of the premium rate. As in Section 7.4 we simulated values of  $\alpha$  and  $\beta$  using a multivariate normal distribution,  $MN(a, b)$  where  $a$  is the mean vector  $(\alpha^{BC}, \beta^{BC}, \alpha^{OC}, \beta^{OC})$  and  $b$  is the covariance matrix.

#### Antoniou *et al.* (2006)

For ‘earliest possible age at onset’ we used:

$$a = (15.56, 0.38, 14.71, 0.36) \text{ and } b = \begin{pmatrix} 0.3300 & 0.0078 & 0.0000 & 0.0000 \\ 0.0078 & 0.0003 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.8500 & 0.0202 \\ 0.0000 & 0.0000 & 0.0202 & 0.0006 \end{pmatrix}$$

and for ‘latest possible age at onset’ we used:



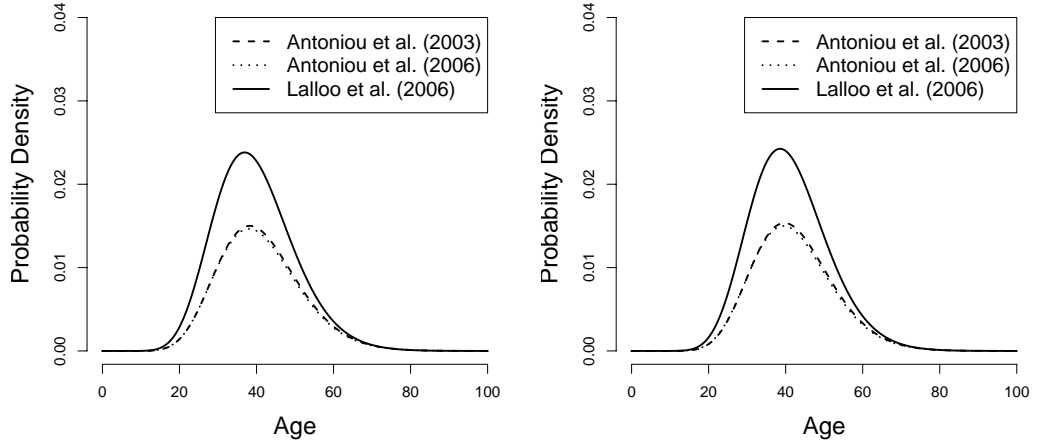


Figure 8.2: Distributions of age at onset for Ovarian Cancer obtained using the penetrance estimates of Antoniou *et al.* (2006), Antoniou *et al.* (2003) and Lalloo *et al.* (2006); Left: Assuming ‘Earliest age at onset’ and Right: Assuming ‘Latest age at onset’.

$$a = (17.34, 0.41, 16.38, 0.39) \text{ and } b = \begin{pmatrix} 0.6700 & 0.0158 & 0.0000 & 0.0000 \\ 0.0158 & 0.0005 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.3100 & 0.0306 \\ 0.0000 & 0.0000 & 0.0306 & 0.0008 \end{pmatrix}$$

### Lalloo *et al.* (2006)

For ‘earliest possible age at onset’ we used:

$$a = (16.36, 0.41, 14.66, 0.37) \text{ and } b = \begin{pmatrix} 0.9300 & 0.0224 & 0.0000 & 0.0000 \\ 0.0224 & 0.0007 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.1100 & 0.0269 \\ 0.0000 & 0.0000 & 0.0269 & 0.0008 \end{pmatrix}$$

and for ‘latest possible age at onset’ we used:

$$a = (18.35, 0.45, 16.42, 0.40) \text{ and } b = \begin{pmatrix} 0.4300 & 0.0102 & 0.0000 & 0.0000 \\ 0.0102 & 0.0004 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.5100 & 0.0122 \\ 0.0000 & 0.0000 & 0.0122 & 0.0004 \end{pmatrix}$$

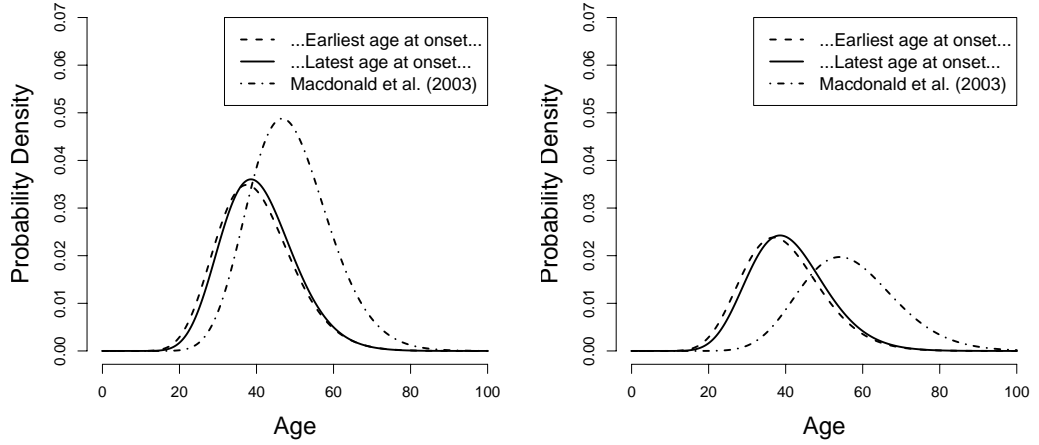


Figure 8.3: Distribution of age at onset for Breast Cancer (left) and Ovarian Cancer (right). Both ‘earliest age at onset’ and ‘latest age at onset’ are shown, assuming the penetrance estimates obtained by Lalloo *et al.* (2006). The distributions used by Macdonald *et al.* (2003) were obtained by fitting truncated Gamma distributions to the estimated incidence rates of Ford *et al.* (1998) for BRCA1 carriers.

Tables 8.5 and 8.6 show the bootstrapped estimates of the premium rate, obtained using the penetrance estimates of Antoniou *et al.* (2006), for the ‘earliest age at onset’ and ‘latest age at onset’, respectively, while Tables 8.7 and 8.8 show the equivalent results using the penetrance estimates of Lalloo *et al.* (2006). We only calculate premium rates for females as male breast cancer is extremely rare in the population.

Based on Tables 8.5 – 8.8 and the chosen combinations of age, policy term and penetrance estimates illustrated in Figures 8.4 and 8.5 we make the following comments:

- (a) The results are affected by the penetrance estimates, and therefore the estimated distribution of age at onset. Using the penetrance estimates of Lalloo *et al.* (2006) results in consistently higher premiums. As the penetrance estimates of Lalloo *et al.* (2006) are higher than that of Antoniou *et al.* (2006) this is to be expected, the higher the penetrance of a disorder the higher the probability of a mutation carrier developing the disorder. An increase in the probability of developing a disorder will result in an increase in the insurance

premiums.

- (b) The assumed age at onset, earliest or latest, does not affect the result significantly. The premium rates would be slightly higher for insurance taken out at ages 20 and 30 if the 'earliest age at onset' was assumed, and higher at ages 40 and 50 if we assumed 'latest age at onset'.
- (c) An applicant known to carry a BRCA1 mutation would be refused insurance for any age or term, regardless of the assumed age at onset for the missing data or the penetrance estimates used as the premium is higher than 350% of the ordinary premium rate. Again, regardless of the assumptions made, the only age at which an applicant would have a possibility of being offered insurance, though unlikely, would be at age 50 for a term of 10 years.
- (d) Although very high at age 20 the standard deviation of the estimated percentage of the OR premium does decrease as the term of policy increases.
- (e) From Figure 8.4 we can see that the percentage of the ordinary premium rate is highest at age 30, for a term of ten years. The ordinary premium rate at age 30 would not be very high as the risk of CI at this age would be low in the general population. The risk of CI for BRCA mutation carriers, however, would be very high at age 30, giving a high% of the OR. The risk at age 20 would be also low in the general population but the excess risk associated with being a mutation carrier would not have come into effect at age 20. At age 40 and 50 the excess risk of mutation carriers would be high but not as high, in comparison to the risk in the general population, as at age 30.

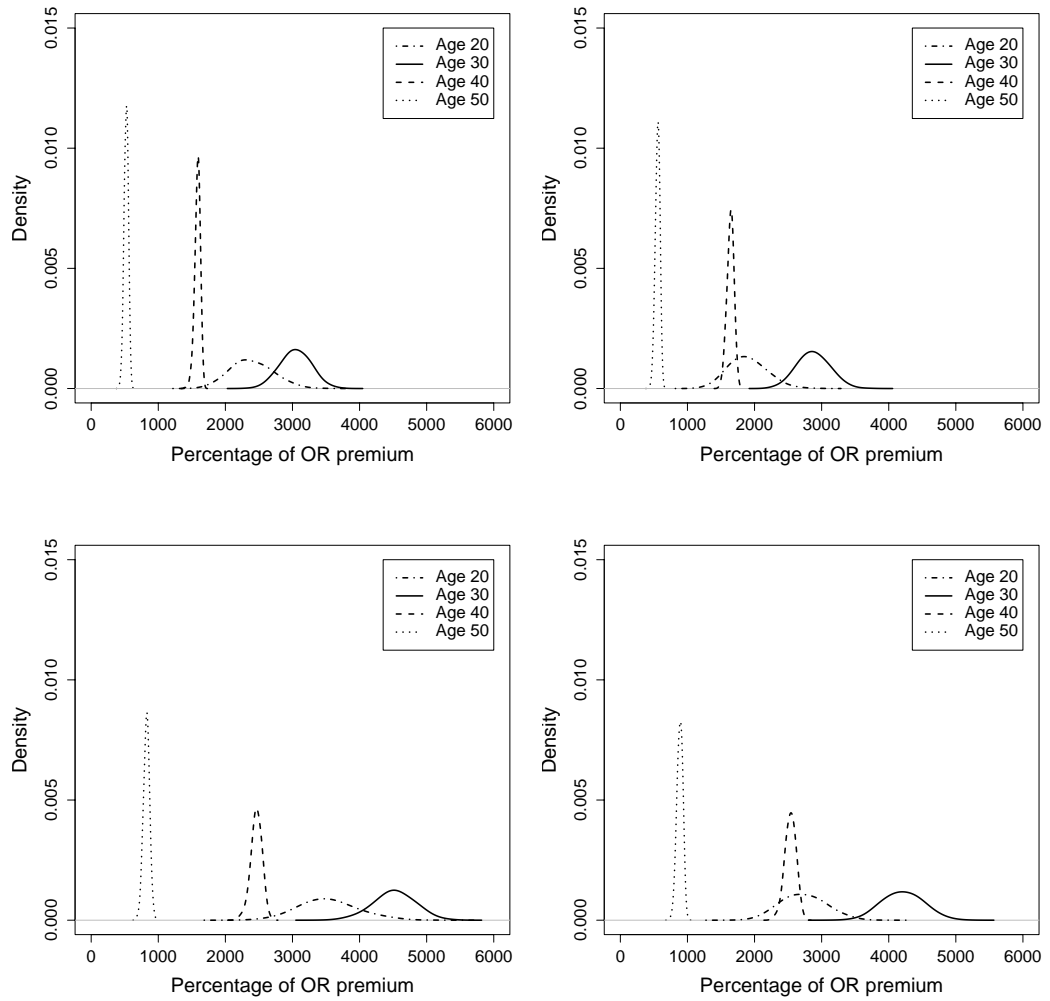


Figure 8.4: Sampling distribution of the premium rate for a range of ages, for an insurance policy term of 10 years.

Top Left: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest age at onset’.

Top Right: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest age at onset’.

Bottom Left: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest age at onset’.

Bottom Right: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest age at onset’.

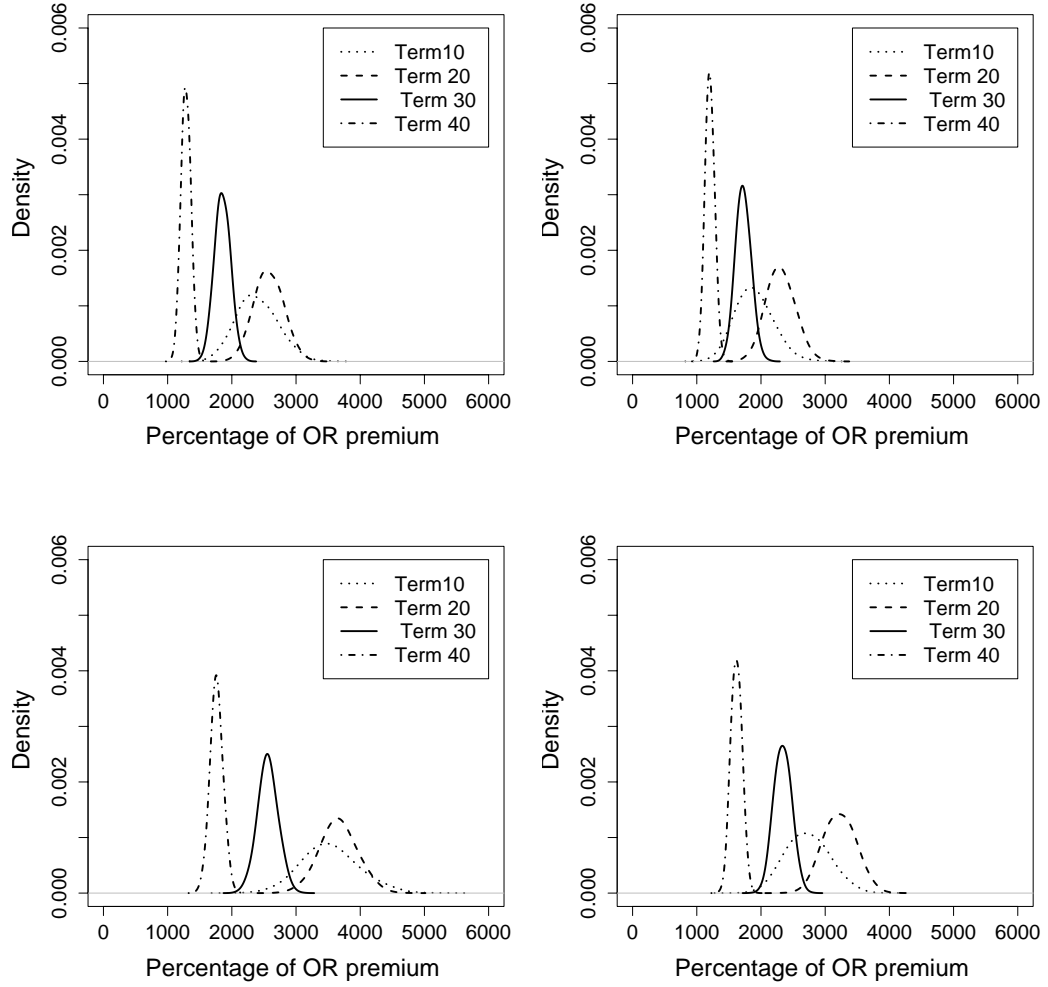


Figure 8.5: Sampling distribution of the premium rate for a range of policy terms, given the age at which insurance was taken out was 20 years old.

Top Left: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest age at onset’.

Top Right: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest age at onset’.

Bottom Left: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest age at onset’.

Bottom Right: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest age at onset’.

Table 8.5: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimates.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,392	301.4
		20	2,576	216.5
		30	1,853	117.0
		40	1,285	71.9
	30	10	3,047	221.9
		20	2,065	111.2
		30	1,400	68.8
		40	1,585	40.5
	50	20	1,022	16.2
		10	522	30.3

Table 8.6: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	1,894	274.9
		20	2,305	212.9
		30	1,724	114.6
		40	1,204	69.8
	30	10	2,877	235.0
		20	2,011	119.1
		30	1,373	68.5
		40	1,640	50.0
	50	20	1,065	20.1
		10	555	32.9

Table 8.7: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Laloo *et al.* (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	3,531	425.6
		20	3,661	279.3
		30	2,560	151.1
		40	1,761	96.5
	30	10	4,515	302.4
		20	3,041	160.2
		30	2,072	98.9
		40	2,462	81.4
	50	20	1,617	40.2
		10	824	43.3



Table 8.8: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,725	382.4
		20	3,225	243.1
		30	2,338	130.9
		40	1,618	82.7
	30	10	4,209	295.8
		20	2,918	157.5
		30	1,999	97.3
		40	2,534	80.9
	50	20	1,679	37.2
		10	886	42.4

## 8.4 Ascertainment Bias

We adjust for ascertainment, as we did for Huntington’s Disease, by randomly removing one proband from each pedigree, as we do not know which pedigree member is the proband and then simulating values from the multivariate normal distribution with the mean being each set of parameter estimates in turn and the variance being the corresponding covariance matrix. However, as discussed in Section 5.9, in the case of breast and ovarian cancer the pedigrees are not usually ascertained through one proband alone. Therefore we also adjust for ascertainment bias by randomly removing two probands from each pedigree and maximising the likelihood using the remaining data. It is of interest to see whether the removal of an extra proband makes any difference to the resulting premium rate estimates.

Tables 8.9 and 8.10 illustrate the premium rate results using the penetrance estimates of Antoniou *et al.* (2006) and randomly removing one proband from each pedigree, for ‘earliest’ and ‘latest’ age at onset, respectively. Tables 8.11 and 8.12 show the corresponding results when two probands were randomly removed from each pedigree.

Tables 8.13 – 8.16 show the same results using the penetrance estimates from Laloo *et al.* (2006).

We make the following comments:

- (a) Adjusting for ascertainment bias, by removing randomly selected probands, does not significantly affect the premium rate.
- (b) The premium rate is, however, higher without any ascertainment adjustment than when any probands are removed for the majority of the policies and terms. As with HD, the exception to this occurs at age 50 when the premium rate is lower when no adjustment is made for ascertainment bias.
- (c) The standard deviation of the premium rate estimates are higher when 2 probands are removed as compared to when only one proband was removed.
- (d) The standard deviation when no proband is removed is consistently lower than when any adjustment is made for ascertainment bias.

- (e) Even after adjustment for ascertainment bias applicants would still be refused insurance at all ages and for all policy terms.

Figures 8.6 - 8.9 illustrate the differences in the sampling distributions of the premium rates for a variety of policy terms, ages and methods of adjustment for ascertainment bias. The average premium rate is reduced, though not significantly when adjusting for ascertainment bias, evident in the plots. The differences in the plots represents the differences in estimated % of the OR premium and differences in the standard deviation of these estimates.

Although an effect of ascertainment adjustment is evident it does not appear to significantly affect the resulting premium rates. One explanation for a lack of effect could be the due to the use of published penetrance estimates in the distributions of age at onset for both disorders, rather than obtaining estimates for the penetrance as two more parameters in the maximum likelihood estimation.

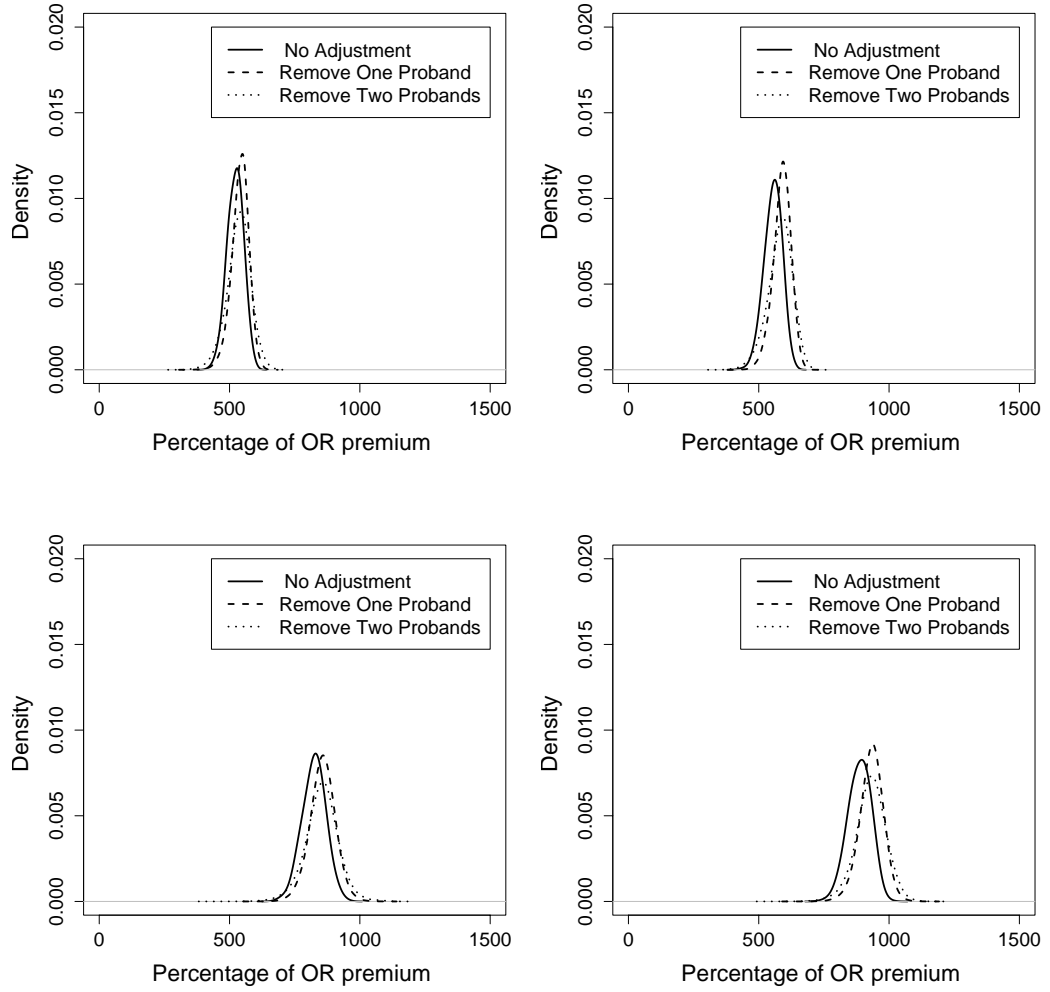


Figure 8.6: Sampling distributions of the premium rate, for a female aged 50 seeking an insurance policy term of 10 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree.

Top Left: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest age at onset’.

Top Right: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest age at onset’.

Bottom Left: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest age at onset’.

Bottom Right: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest age at onset’.

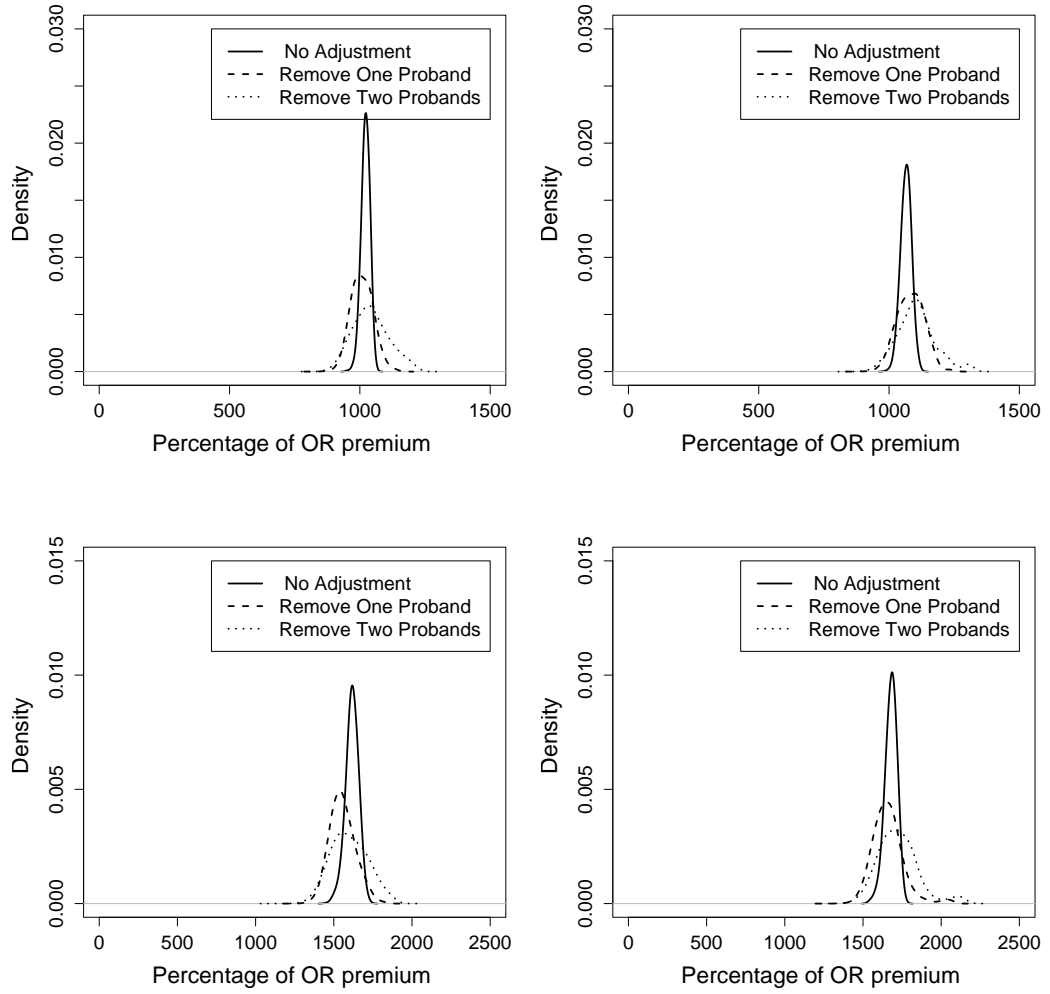


Figure 8.7: Sampling distribution of premium rate, for a female aged 40 seeking an insurance policy term of 20 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest age at onset’.

Top Right: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest age at onset’.

Bottom Left: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest age at onset’.

Bottom Right: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest age at onset’.

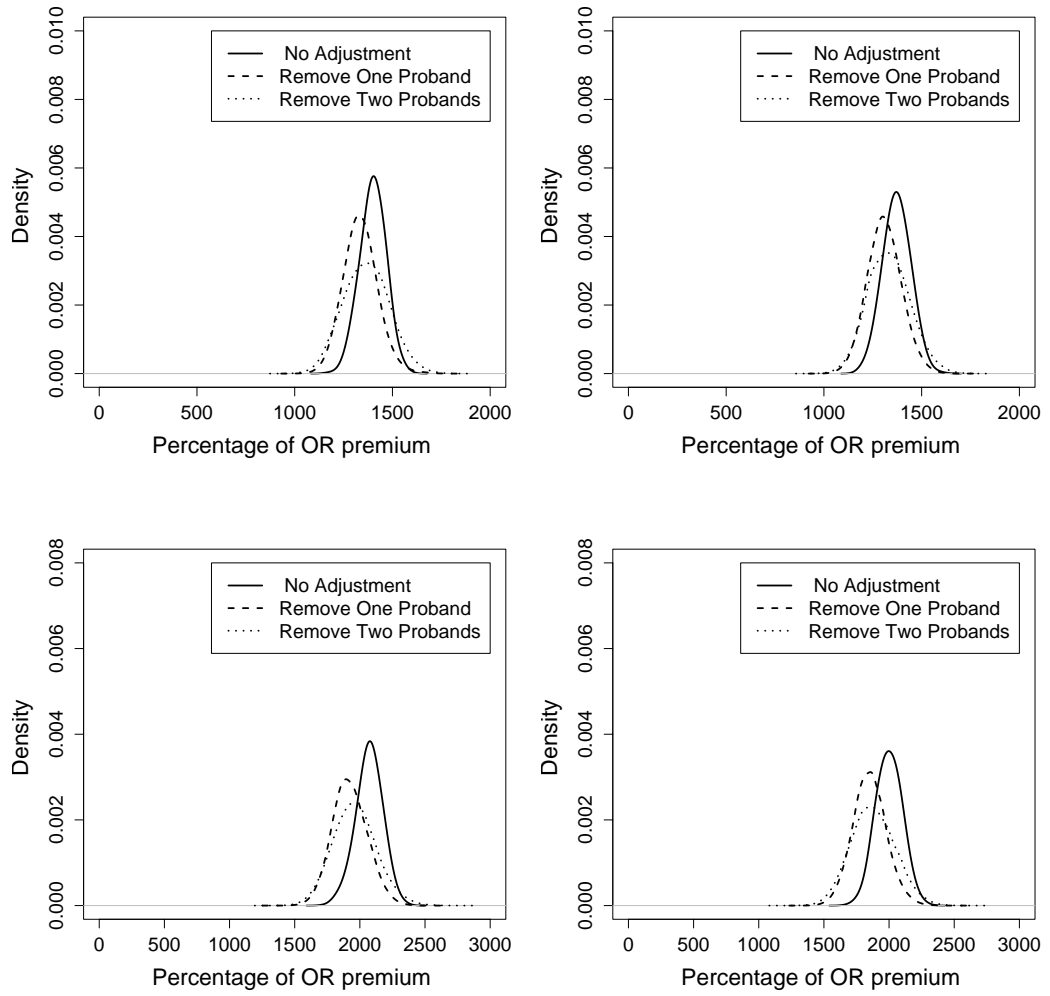


Figure 8.8: Sampling distribution of premium rate, for a female aged 30 seeking an insurance policy term of 30 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest age at onset’.

Top Right: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest age at onset’.

Bottom Left: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest age at onset’.

Bottom Right: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest age at onset’.

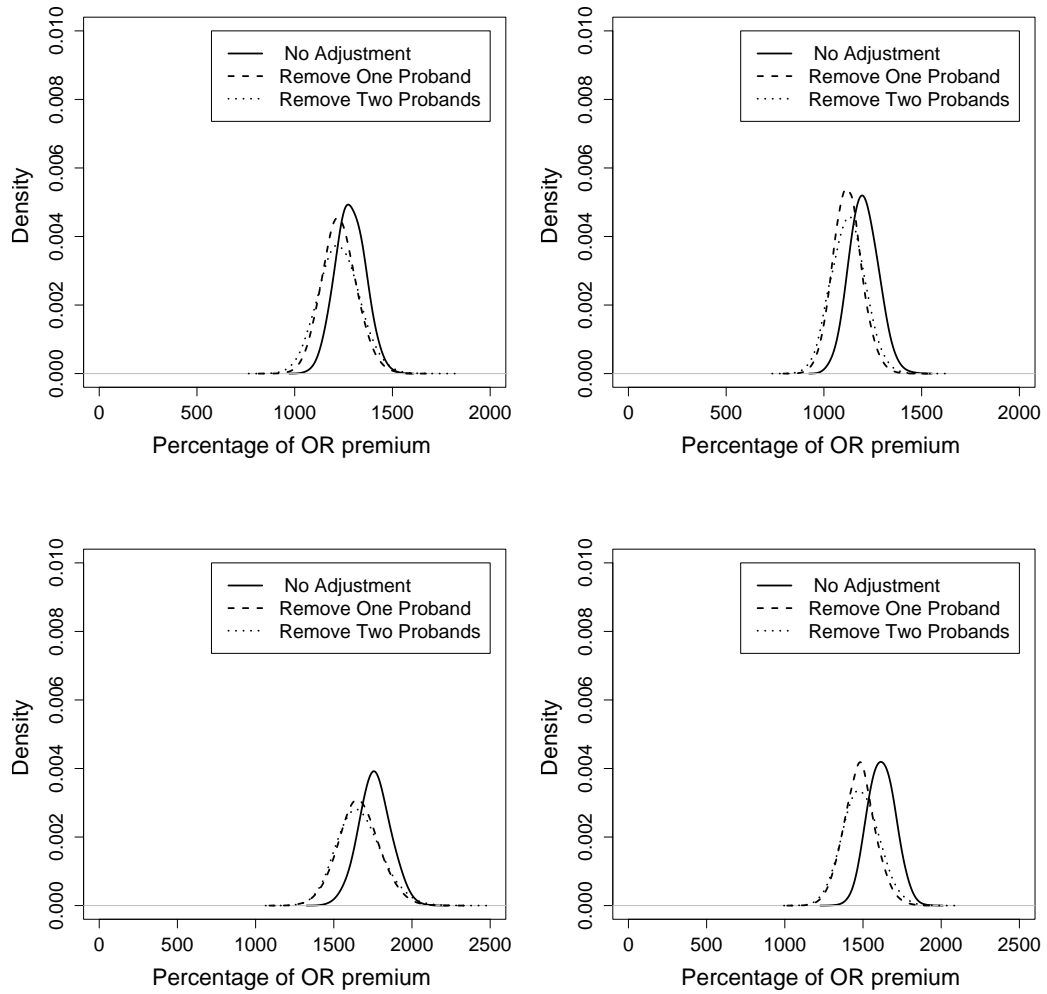


Figure 8.9: Sampling distribution of premium rate, for a female aged 20 seeking an insurance policy term of 40 years with; (a) no adjustment for ascertainment bias, (b) adjusting for ascertainment bias by removing one proband from each pedigree and (c) adjusting for ascertainment bias by removing two probands from each pedigree. Top Left: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest age at onset’.

Top Right: Using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest age at onset’.

Bottom Left: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest age at onset’.

Bottom Right: Using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest age at onset’.

Table 8.9: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,201	389.2
		20	2,424	267.3
		30	1,755	143.0
		40	1,228	89.6
	30	10	2,850	278.5
		20	1,974	158.1
		30	1,336	88.5
		40	1,537	90.2
	50	20	1,008	43.4
		10	540	33.8



Table 8.10: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	1,487	296.8
		20	2,032	225.8
		30	1,586	122.7
		40	1,120	74.4
	30	10	2,633	268.9
		20	1,906	148.0
		30	1,314	86.7
		40	1,645	107.7
	50	20	1,083	52.6
		10	589	34.6

Table 8.11: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,148	502.0
		20	2,451	197.6
		30	1,768	171.4
		40	1,227	105.4
	30	10	2,909	363.4
		20	2,052	197.6
		30	1,366	116.4
		40	1,579	133.4
	50	20	1,042	69.8
		10	535	46.2

Table 8.12: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Antoniou *et al.* (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	1,445	365.7
		20	2,042	267.1
		30	1,598	143.2
		40	1,125	87.2
	30	10	2,687	338.7
		20	1,946	188.7
		30	1,333	108.4
		40	1,693	154.2
	50	20	1,105	74.3
		10	581	47.8

Table 8.13: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	3,186	597.9
		20	3,349	326.1
		30	2,383	199.9
		40	1,662	131.7
	30	10	4,096	383.2
		20	2,812	216.9
		30	1,924	134.9
		40	2,324	153.8
	50	20	1,555	83.2
		10	857	49.7

Table 8.14: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing one proband.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,240	440.8
		20	2,833	295.7
		30	2,124	160.0
		40	1,478	100.9
	30	10	3,751	359.1
		20	2,703	211.5
		30	1,851	129.3
		40	2,426	175.3
	50	20	1,659	100.4
		10	933	46.5

Table 8.15: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘earliest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	3,131	720.3
		20	3,353	431.7
		30	2,395	231.3
		40	1,665	142.6
	30	10	4,153	490.6
		20	2,878	274.8
		30	1,961	167.4
		40	2,373	203.6
	50	20	1,600	118.9
		10	853	64.8

Table 8.16: Level net premium for level CI cover for persons with a known BRCA1 mutation, as a percentage of the premium for the OR class. Using parameters obtained given assumption A1, using the penetrance estimates of Lalloo *et al.* (2006) and assuming ‘latest possible age at onset’. The table also includes the standard deviation of the estimated premium. These are the results obtained after adjusting for ascertainment bias by removing two probands.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	% of OR Premium	Standard Deviation of Premium
Female	20	10	2,178	525.8
		20	2,839	357.8
		30	2,121	186.2
		40	1,488	115.8
	30	10	3,781	443.3
		20	2,781	252.6
		30	1,878	168.0
		40	2,476	242.2
	50	20	1,724	130.4
		10	930	59.2

## 8.5 Conclusions

Based on the distributions of the age at onset for breast and ovarian cancer, for BRCA1 carriers, estimated using the pedigrees available to us, it would be very unlikely for a female to be offered critical illness insurance at any age or for any policy term.

As would be expected the estimate of the penetrance used affects the estimated distribution of age at onset and the resulting premium rate sampling distributions; using higher penetrance estimates results in higher premium rates. As critical illness insurance would not be offered at any age or term, however, it does not make any difference to the conclusions reached in this investigation although it is something that should be thought about in any future studies. Using reliable penetrance estimates when estimating the distribution of the age at onset will result in more reliable premium rate estimates.

As with Huntington's Disease the choice of whether to use 'earliest' or 'latest' age at onset for those whose age at onset was missing does not make any real difference to the resulting premium rates.

Adjustment for ascertainment bias, either by removing one or two randomly selected probands from each pedigree, increases the standard deviation of the estimated premium and reduces the mean premium rate slightly. This is an interesting result illustrating that, although there is some change in the sampling distribution of the premium rates depending on whether ascertainment bias is adjusted for or not, it is small and does not result in any change in the conclusions reached about whether a person would be insured given that they were known to be a BRCA1 mutation carrier. It would be interesting to see if this were still the case if we estimated the penetrance of breast and ovarian cancer rather than using published estimates.

As with HD there is evidence that the premium rates that a BRCA1 mutation carrier would be charged would result in them being declined critical illness insurance for most ages and policy terms. This suggests that there is evidence that the genetic test for BRCA1 has actuarial significance.



## Chapter 9

# Familial Adenomatous Polyposis

### 9.1 Data

We have data from 40 FAP pedigrees supplied by Cardiff University, including 152 persons with FAP. Some had to be excluded because of lack of information about year of birth, leaving 132 with known age at onset. Of these, 107 were diagnosed by screening, and 25 were not identified as having FAP until they were diagnosed with CRC or until they died.

The discussion of the analysis of the FAP data deserves a chapter to itself, as the FAP data and FAP disease take on a slightly different form than the other disorders considered.

### 9.2 How Might Screening Affect Insurance Risk

Here we are interested in how the availability of screening, and now genetic testing, for FAP may affect life and critical illness insurance. We ask: (a) does screening reduce the risk of CRC and death? and (b) does better targetted screening based on genetic tests reduce that risk yet further?

For now insurers are able to make use only of family history of FAP when determining premiums. A person with a family (parental) history of FAP has a 50% chance of also developing FAP and this is reflected in the premium calculation. However if

insurers were able to use the results of genetic tests, a person known to be at high risk of developing FAP, following a genetic test, would be likely to have a higher premium.

If genetic tests can be used by insurers, people with a family history of FAP may be unwilling to be genetically tested and screened as these results would have to be disclosed and would be detrimental to their insurance premium. On the other hand people may get tested and screened but not attempt to buy insurance as they believe that they would not be successful or that their premiums would be too high for them.

Here we aim to show that in the case of FAP the use of genetic testing and screening can be very beneficial to possible FAP sufferers and therefore there may be no need for insurers to consider FAP to be a problem when calculating premiums, as long as the possible sufferers get genetically tested and follow a regular screening program.

Ideally, we would carry out a standard survival analysis to estimate rates of onset of FAP. However, this is ruled out because we have almost no information about those who did not suffer onset, that is, censored observations. Therefore we are limited to analysing those whose age at onset was observed, for the same reasons we are also limited in the analysis that can be carried out. Although this ‘sample’ is obtained retrospectively, we may legitimately ask questions about its composition conditioning on the fact of onset. We will look at three factors: calendar year of onset; year of birth; and method of detecting FAP.

### 9.3 Year of Onset of FAP

Affected persons were divided into two groups, those who had suffered onset before a given calendar year  $Y$  and those who had suffered onset in or after year  $Y$ . Figure 9.1 shows the empirical survival functions (‘survival’ meaning free of FAP) for these two groups, on the left with  $Y = 1990$ , and on the right with  $Y = 2000$ . The former shows little difference between the two groups; the latter suggests there is a slight difference but only 13 people suffered onset after 2000. Formally, using the Wilcoxon Mann-Whitney (WMW) test of equal medians, there is no difference in

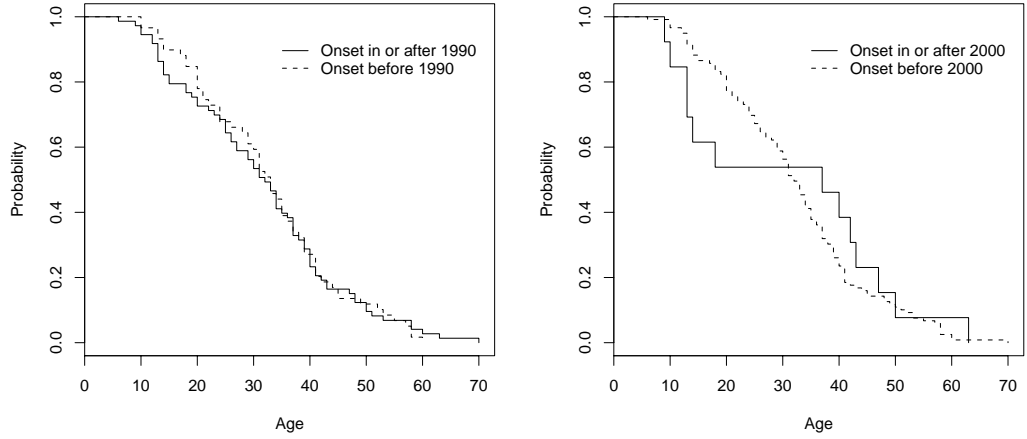


Figure 9.1: Left: Probability of survival given ‘onset occurred before 1990’ and probability of survival given ‘onset in or after 1990’.  
Right: Probability of survival given ‘onset occurred before 2000’ and probability of survival given ‘onset in or after 2000’.

either case ( $Y = 1990$ ,  $p = 0.616$  and  $Y = 2000$ ,  $p = 0.798$ ).

## 9.4 Year of Birth

The earliest year of birth in the data is about 1900 and the latest about 2000. We compared two groups, those born before calendar year  $B$  and those born in or after year  $B$ . Figure 9.2 shows, on the left, empirical survival functions for  $B = 1950$ , and on the right for  $B = 1970$ . There is a clear difference but the maximum possible age at onset is different for each group; we are not comparing like with like. Specifically, for each choice of  $B$  we are comparing the random variables  $X_1$  and  $X_2$  where:

$$X_1 = \text{Age at onset given Born before year } B \quad (9.1)$$

$$X_2 = \text{Age at onset given Born in or after year } B. \quad (9.2)$$

However, since we are viewing these data retrospectively from the present time  $T$ ,  $X_1$  and  $X_2$  have different ranges:  $X_1$  ranges over all human lifetimes whereas  $X_2$  cannot exceed  $T - B$ . To compare like with like, we must exclude those born before

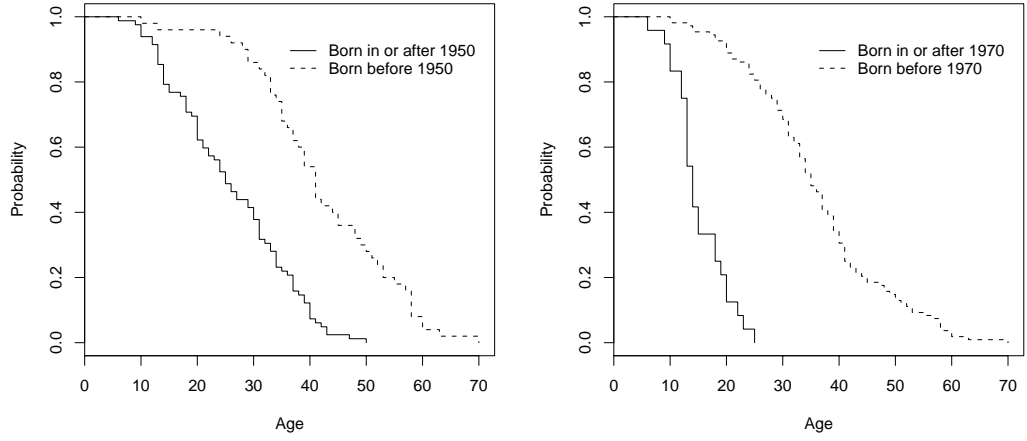


Figure 9.2: Left: Probability of survival given ‘born before 1950’ against the probability of survival given ‘born in or after 1950’.  
Right: probability of survival given ‘born before 1970’ against the probability of survival given ‘born in or after 1970’.

year  $B$  whose age at onset exceeds  $T - B$ . In other words, compare  $X_1^*$  and  $X_2^*$  where:

$$X_1^* = \text{Age at onset given Born before } B \text{ \& not exceeding } T - B \quad (9.3)$$

$$X_2^* = \text{Age at onset given Born in or after year } B. \quad (9.4)$$

Figure 9.3 shows the results with  $X_1^*$  limited to age 50 ( $B = 1950$ ) and age 30 ( $B = 1970$ ). A difference remains in both cases. This suggests that those born in or after 1950 were diagnosed earlier than those born before 1950

## 9.5 Method of Diagnosis

Finally, we stratify by the method of diagnosis, whether by screening or in some other way (death, onset of cancer and so on). Figure 9.4 suggests that the curves are different until they cross at nearly age 60. The hypothesis test of equal medians shows a difference ( $p < 0.001$ ). Screening does seem to result in an earlier diagnosis of FAP and reduced cancer risk.

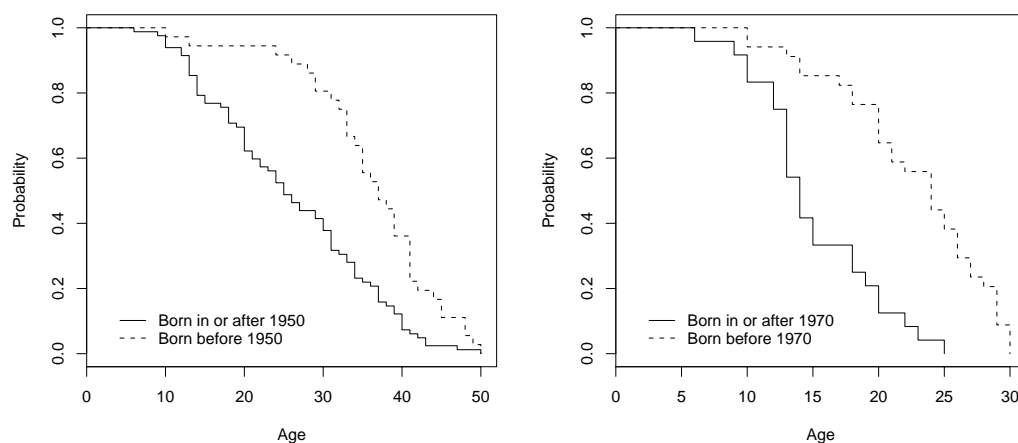


Figure 9.3: Left: Probability of survival given ‘born before 1950 and onset before age 50’ against the probability of survival given ‘born in or after 1950 and onset before age 50’.

Right: Probability of survival given ‘born before 1970 and onset before age 30’ against the probability of survival given ‘born in or after 1970 and onset before age 30’.

## 9.6 Screening and Cancer Risk

Figure 9.5 illustrates how screening and now genetic testing have altered substantially the cancer risk associated with FAP. It presents three alternative life histories. At the top is the most likely life history in the absence of screening: the first event is the diagnosis of CRC followed by post-onset diagnosis of FAP. In the middle is the life history of someone who undergoes screening because of a family history of FAP. The screening allows treatment to follow the diagnosis of FAP, with fewer cases proceeding to CRC. At the bottom is the identification of high-risk individuals by genetic testing, making screening more effective.

In our sample, 152 people fall into the three groups; 26 suffered CRC without any previous diagnosis of FAP (top), 90 were diagnosed with FAP through family screening (middle), and 36 were determined to be at high risk through genetic testing (bottom). (Note that year of birth was not needed here so more individuals could be included here than above.)

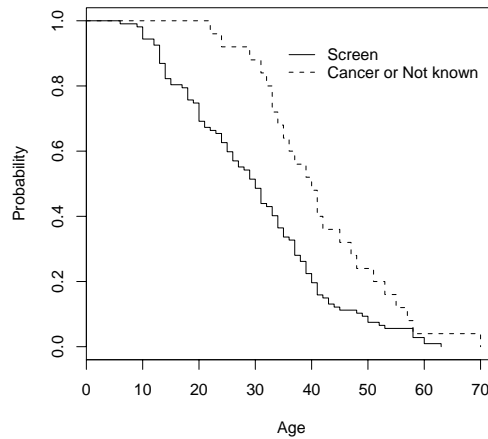


Figure 9.4: Time to onset for those diagnosed through screening compared with those diagnosed through other methods.

Of the 152 people who have FAP 17.1% developed CRC without any diagnosis of FAP, 6.6% developed CRC after being diagnosed through family screening while no-one so far has developed CRC after genetic testing. Of course, genetic testing is relatively recent. the latter may be because it is too soon for these people to have developed CRC, or because their risk of suffering CRC after early treatment is genuinely very low.

Screening and genetic testing does enable earlier diagnosis. The average ages at diagnosis of those diagnosed through having CRC, those diagnosed through screening and those diagnosed through screening as a result of genetic testing is 42, 32 and 19 respectively. Since most people have surgery as soon as FAP is diagnosed (88% of people in the pedigrees used here) this also reduces the risk of developing CRC. Järvinen (1992) determined that those diagnosed through symptoms alone have a 65.5% risk of colorectal cancer, whereas those diagnosed through family screening have a 6.6% risk. This is only slightly higher than the population lifetime risk of CRC, 2 – 5%. Evans (1993) says that screening at risk patients with a combination of genetic markers, screening for congenital hypertrophy of the retinal pigment epithelium and a negative bowel investigation may reduce initial risk of cancer from 50% to well below 1%.

Both Bülow (2003) and Heiskanen (2000) say that the use of screening improves

the survival of FAP sufferers. Bülow found that the survival rate is 44% for people diagnosed with FAP through symptoms (probands) and 94% for those diagnosed through screening based on family history (call-up). Heiskanen found that there was a significant increase in survival for the call-up patients compared to the probands after colectomy or proctocolectomy.

## 9.7 Conclusions

Here we address the questions: (a) does a family history of FAP, or a genetic test result, indicate significantly increased life and critical illness insurance risk? and (b) is there any reason for a person in a high-risk family to be deterred from undergoing testing and screening because of the insurance implications?

The results above show that the age of diagnosis of FAP is even lower when screening follows genetic testing than when screening is used alone. As the risk of CRC is reduced to almost the population level (3 – 5%; Evans, 1993) by screening alone (6.6%; Järvinen, 1992) it is plausible that genetic testing will reduce the risk further; how much further we cannot yet say. It may be that earlier treatment, through better targetted screening, will eliminate the excess risk.

The risk of CRC need no longer be significantly higher for those at risk of FAP, hence insurance premiums need not be significantly higher either, provided the individual is actively following a screening program. Those at risk should not be deterred from being tested and screened because of any perceived implications for life and critical illness insurance. We suggest that, in this case, insurance and clinical interests are aligned instead of in opposition. However, this leaves open the question of how an insurer should respond in case a person exercises their autonomy over medical decisions, and chooses not to be tested or screened.

The findings here suggest that co-operation between those with a family history of FAP and insurers would result in benefits for both parties. FAP sufferers could get tested, screened and treated without worrying about how it will affect their insurance and insurers can cross FAP off their list of genetic disorders to worry about.

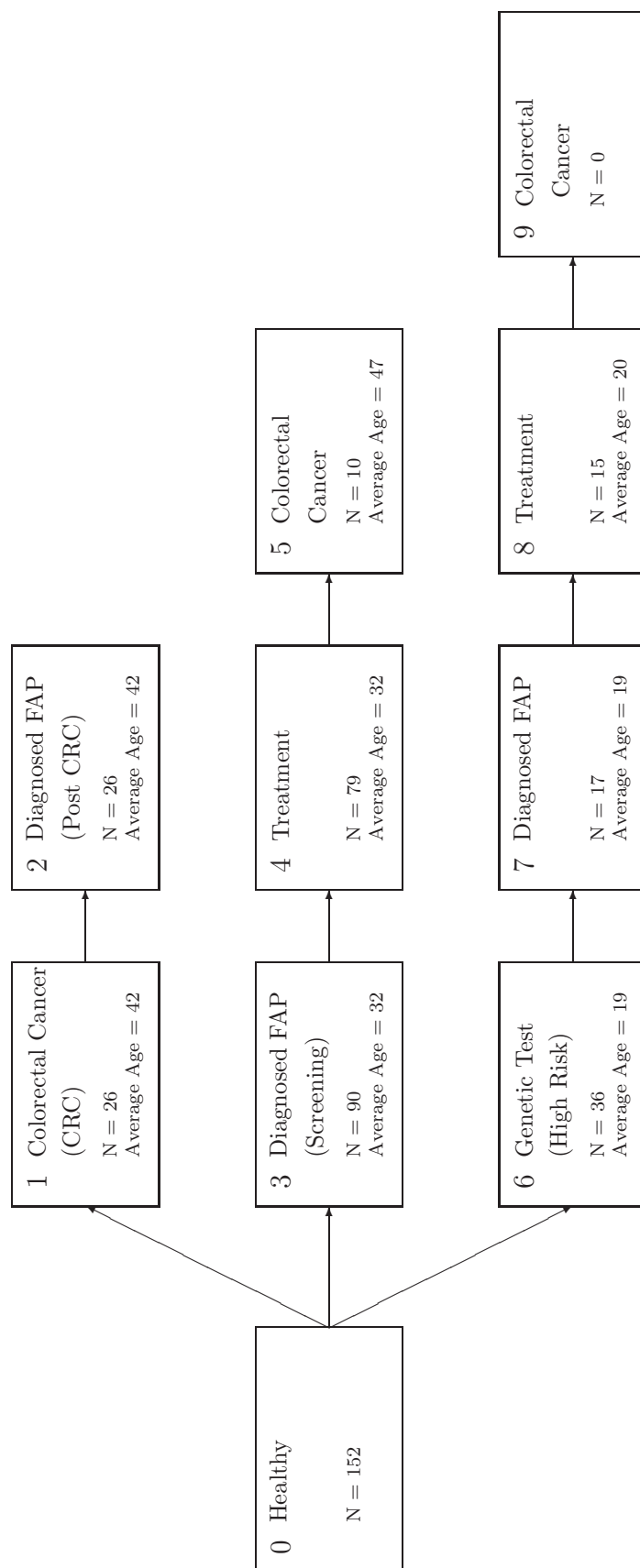


Figure 9.5: States of those who have been confirmed to have FAP, through testing, diagnosis (screening) and colorectal cancer.  $N$  is the number of people that have been in each state. The average age is given as the average age, on entering the state, of all people who have been in this state.



# Chapter 10

## Conclusions and Recommendations for Future Work

In this investigation we have determined the premium rates that would be charged to known mutation carriers for two specific genetic disorders, Huntington's Disease and BRCA1-related breast and ovarian cancer, in order to see the possible effects that knowing the result of a genetic test for either of these disorders would have. However, as we had pedigree data available for both of these genetic disorders, we were able to do more than just determine the premium rate appropriate for a known mutation carrier.

Very few studies concerning genetic disorders and insurance have been able to determine the sampling distribution of the premium rate as neither the variability of the parameters of the distribution of the age at onset nor the data behind these estimates are routinely available. Here we used the available pedigree data to determine the variance matrix for the parameters of the distribution of age at onset. As a result we were then able to estimate not only the premium rate but also the variance, and therefore a measure of reliability, of the premium rate.

Another area of interest when examining genetic disorders is the effect of ascertainment bias on the distribution of the age at onset and the resulting premium rates. Again, as we had the pedigree data available, we were able to adjust for

ascertainment bias while estimating the distribution of age at onset and determine the sampling distribution of the premium rate after adjustment for ascertainment bias was carried out.

Being able to use actual pedigree data allows the investigation of genetic disorders and the premium rates associated with them to move a step further.

## **10.1 Huntington's Disease**

### **10.1.1 Onset Rates**

The distribution of age at onset for HD was assumed to follow a Normal distribution. The estimated parameters of the distribution of the age at onset were higher than those estimated by the majority of previous epidemiological studies (see Table 4.2). A plot of the estimated distributions of age at onset and the distribution of age at onset of two published studies showed that the differences were relatively small (see Figure 7.2).

The different assumptions made about the data missing from the pedigrees made little overall impact on the parameter estimates. The decision of whether to assume 'earliest' or 'latest age at onset' resulted in small, insignificant differences in the estimates.

### **10.1.2 Premium Rates**

As HD is a fully penetrant genetic disorder it is unlikely for a mutation carrier to not suffer onset of the disorder if they live long enough. This, coupled with the low mean age at onset, obtained from the pedigree data, of about 44 years old, results in there being very few instances where a known HD mutation carrier would be offered critical illness insurance.

The definition of CI event used by insurers, represented here by Stage 2 or Stage 3 of Harper's progression, makes little difference to whether a known mutation carrier would be offered insurance or not, with the only exception being that using Stage 3

as the CI event would result in mutation carriers being able to take out insurance at age 50 for a term of 10 years.

The estimated sampling distribution of the premium rates show that there is more variance in the estimated premium rate at lower ages and policy terms. Any confidence intervals, based on the sampling distributions, would not include 500% of the standard premium, with the exception of a policy taken out at age 50 for 10 years given the CI event being Stage 3 of Harper’s progression. It is also clear that even charging the minimum bootstrapped premium rate would result in an insurance application being denied.

There is sufficient actuarial evidence, in the case of CI insurance, that the premium rates are increased for known mutation carriers to such a level that they would not be offered insurance. This, in itself, is not new but the methodology that we have used here, based on data from real pedigrees, parametric modelling and bootstrapping provides a robust criterion for any approach that GAIC might take to applications in future, which in general may not be as clear-cut as that of HD (see Section 10.4).

## **10.2 BRCA1 related breast and ovarian cancer**

### **10.2.1 Onset Rates**

For BRCA1-related breast and ovarian cancer it was assumed that the distributions of age at onset followed Gamma distributions. The parameters that were estimated using the pedigree data, and assuming the penetrance to be that estimated in a selection of previous studies, were very similar for both breast and ovarian cancer. As with HD, the different methods used to deal with the missing data in the pedigrees did not have a significant effect on the resulting parameter estimates, with the exception that whether the ‘earliest’ or ‘latest age at onset’ was used makes a small difference to the estimated parameters, but no significant difference to the actual distribution of age at onset.

The penetrance estimates used did not significantly affect the estimates of the pa-

rameters, but made a difference to the overall age at onset distribution (see Figures 8.1 and 8.2). The estimated distribution of the age at onset for both breast and ovarian cancer are very similar to that used by Macdonald *et al.* (2003a), see Figure 8.3.

### 10.2.2 Premium Rates

BRCA1-related breast and ovarian cancer are highly penetrant disorders so it is therefore not a surprise that the premium rates that would be charged to a known BRCA1 mutation carrier would mean that they would not be offered insurance for the majority of ages and policy terms.

The choice of penetrance estimates used when estimating the distributions of age at onset has an effect on the sampling distributions of the premium rates, with the rates resulting from using the higher penetrance estimates of Lalloo *et al.* (2006) being higher than when using the penetrance estimates of Antoniou *et al.* (2006). However, due to the high penetrance of both breast and ovarian cancer, regardless of which penetrance estimates were used, the resulting sampling distribution of the premium rates would not affect the overall conclusions.

Again the investigation carried out here suggests that the result of a genetic test for BRCA1 would be of interest to insurers as the premium rates that would be appropriate for a known mutation carrier would result in them not being offered critical illness insurance. The standard deviations of the premium rates add to the evidence, in that although there is a lot of variability in the estimated premium rates, especially for lower ages and policy terms, the variability is not so high that an applicant would be offered insurance at a premium rate in the plausible range.

## 10.3 Effect of adjusting for ascertainment bias

Adjustment was made for ascertainment bias by removing a proband from each pedigree, where the proband was randomly selected from the possible probands in each pedigree. In addition to this, for breast and ovarian cancer, an adjustment for ascertainment bias was also made by removing two randomly selected probands

from each pedigree as BRCA1-related breast and ovarian cancer families are likely to have been identified due to multiple cases.

Carrying out adjustment for ascertainment bias for BRCA1-related breast and ovarian cancer resulted in slight differences to the distribution of the premium rates, specifically removing one proband increased the standard deviation and removing two increased it further. As noted in Section 8.4 the lack of a significant effect of ascertainment bias on the location of the sampling distribution of the premium rate for BRCA1-related breast and ovarian cancer could be due to the fact that, instead of estimating the penetrance of both disorders within the likelihood maximisation, we used published penetrance estimates.

In contrast, although the adjustment for ascertainment bias for HD did not make any difference to the overall conclusions concerning whether an applicant would be insurable or not, there was a significant difference in the estimated sampling distributions of the premium rate after adjusting for ascertainment bias. The premium rate was reduced significantly after adjusting for ascertainment bias and the standard deviation was increased. As HD is fully penetrant, and is an extreme example of a genetic disorder, it is not surprising that ascertainment adjustment does not alter the fact that a known mutation carrier would be declined critical illness insurance.

The difference in the sampling distributions of the premium rate for known HD mutation carriers with and without ascertainment bias adjustment highlights the importance of adjusting for ascertainment bias and the effect it has on the premium rate. For a less penetrant disorder the adjustment for ascertainment bias may significantly affect the decision on whether to sell a known mutation carrier insurance or not.

## 10.4 Implications for GAIC criteria

As discussed in Section 3.2 it is up to GAIC to decide whether a genetic test result can be used by an insurer. Applications are made to GAIC, by the ABI, in relation to a specific genetic test and a specific insurance product. Each application must

include:

- 1 Evidence that the genetic test is accurate and reliable.
- 2 Clinical evidence that the the presence of a mutation leads to implications for the health of an individual.
- 3 Actuarial evidence that there is an increased probability of the insured event occurring if the mutation is present.

Evidence in support of both points 1 and 2 can be obtained from epidemiological and genetic studies. The investigations carried out here relate to possible evidence for point 3 of GAIC's criteria list, for critical illness insurance.

It has been shown here, for both Huntington's Disease and BRCA1-related breast and ovarian cancer, that the probability of developing the critical illness event is greatly increased when the related mutation is known to be present. In fact, for both disorders, the probability is increased to such an extent that the premium rates appropriate for a known mutation carrier would result in them being declined critical illness insurance at almost all ages and policy terms.

The results obtained in this thesis confirm the already well-known fact that for both disorders known mutation carriers would not not be offered critical illness insurance. For the first time though we are able to provide an estimate of the reliability of the conclusion, using the sampling distributions of the premium rates. Being able to assess the reliability of the premium rate estimates will be of even more importance when considering disorders that do not have such extremely high penetrance.

## 10.5 Familial Adenomatous Polyposis

Examination of the FAP pedigrees illustrated that genetic testing, in addition to screening, results in possible FAP mutation carriers being diagnosed earlier. The risk of colo-rectal cancer is reduced to almost the population level by screening alone so it is plausible that genetic testing will reduce the risk further

We would conclude that a known FAP mutation carrier should not be put off seeking critical illness insurance. Rather we would say that co-operation between those with a family history of FAP and insurers would result in benefits for both parties. Those with a family history of FAP could then undergo genetic testing, follow the resulting screening program and be treated, in the event that FAP develops, in such a way that the risk of developing colo-rectal cancer as a result of being an FAP carrier would be minimised, or perhaps even eliminated. Insurers would then be able to remove FAP from the list of genetic disorders to worry about in terms of insurance.

## **10.6 Further Work**

For both breast and ovarian cancer and Huntington's Disease the estimation of the distributions of age at onset would benefit from using data from more pedigrees, if it was available.

### **10.6.1 Breast and Ovarian Cancer**

Although we found evidence of an effect of ascertainment adjustment on the estimated sampling distribution of the premium rate it would be interesting to see if this was affected by the fact that we used published penetrance estimates. It would be of interest to carry out the parameter estimation for BRCA1-related breast and ovarian cancer again, but to also estimate of the penetrance of both breast and ovarian cancer based on the data.

By estimating the penetrance estimates, instead of using published estimates, the effect of ascertainment bias adjustment could be investigated in more depth. In particular if the estimation of the penetrance of breast and ovarian cancer resulted in there being an effect of either method of ascertainment bias adjustment then the relationship between the adjustment used and the resulting sampling distributions of the premium rate could then be looked at.

The likelihood maximisation could be extended to include the penetrance of both breast and ovarian cancer as unknown parameters. This would result in a total of six parameters being estimated which would substantially increase computation

time. The increase in the number of the parameters being estimated is also likely to result in more variation in the estimates, unless more data is used.

The distributions of age at onset for breast and ovarian cancer were only estimated for BRCA1 mutation carriers as there were not enough pedigrees available for BRCA2 carriers. Identifying more BRCA pedigrees, especially those families carrying BRCA2, would enable the estimation of the distributions of age at onset of both breast and ovarian cancer for both BRCA1 and BRCA2 mutation carriers. The premium rates could then be compared for BRCA1 and BRCA2 carriers and it could be determined whether there was a significant difference in the effect of BRCA1 and BRCA2 and also to investigate whether the sampling distributions of the premium rates after adjusting for ascertainment bias behaved differently for the two mutations.

Another possibility for future work in the case of BRCA1/2-related breast and ovarian cancer is to consider the model of Antoniou *et al.* (2002). They modelled the effects of BRCA1, BRCA2 and a polygenic effect, finding that their polygene accounted for a large proportion of familial breast cancer. Using the pedigree data available here we could attempt to obtain the sampling distributions after allowing for the polygene effect.

### 10.6.2 Huntington's Disease

If it was possible to determine the CAG repeat lengths for each known mutation carrier within the pedigrees then it could be possible to incorporate this CAG repeat length into the estimation of the age at onset distributions and any subsequent estimation of the premium rate. It would be very difficult and time-consuming to get this data, however, and to be able to estimate the age at onset distribution for a number of CAG repeat lengths it is highly likely that significantly more pedigrees would be required.

In this investigation we adjusted for ascertainment bias by removing randomly selected probands from each pedigree; this was done as we did not know who the actual proband was in each pedigree. With more investigation into the medical records it may be possible to find out which member of each pedigree was actually



the proband. It may also be possible to investigate the effect that different methods of ascertainment adjustment would have on the premium rates and determine how the method used here compares.

### **10.6.3 Familial Adenomatous Polyposis**

As noted in Section 9.2, due to a lack of information about the people who had not suffered onset of FAP, we could not carry out any parametric testing. Although it may be of interest if more information could be found about all members of the pedigrees, perhaps through further investigation of the records or perhaps contact with the families, it would not enable us to investigate FAP in any more depth than has been done here.

As the risk of an FAP mutation carrier developing colo-rectal cancer (CRC), the critical illness event associated with FAP, appears to be decreasing, especially when screening is carried out, any investigation of the age at onset of CRC in FAP carriers would require data from significantly more pedigrees than are available here.

Also as a result of the changes in detection and treatment, the younger generations of families known to carry the FAP mutation are diagnosed, and subsequently treated, earlier. This could result in being unable to treat the older and younger generations of the pedigrees as the same when carrying out modelling and estimation.

# Appendix A

## Adult Polycystic Kidney Disease

### A.1 Epidemiological Review

Adult polycystic kidney disease (APKD) is a genetic disorder that primarily affects the kidneys. Cysts develop in the kidneys causing them to enlarge, often leading to kidney failure. The first symptom of APKD is often high blood pressure as the cysts squeeze the blood vessels. Many other symptoms may follow including fatigue, headaches, frequent urination, blood in the urine, kidney stones and urinary tract infections (Milutinovic *et al.*, 1984).

Serious problems can occur as a result of these symptoms. These include brain aneurysms, enlargement of the heart, cysts in the pancreas and liver as well end-stage renal disease (ESRD). According to the PKD Foundation over 60% of people with APKD develop ESRD, where the kidneys can no longer function and can only be treated through renal replacement therapy (RRT), which consists of dialysis and/or kidney transplantation.

Robinson & Hawkins (1981) estimated that APKD accounts for 6 – 9% of ESRD in Europe, while Davies *et al.* (1991) estimated that this proportion was slightly higher in Wales, 10.4%. In a study carried out in the US it was found that 6.5% of all transplants following ESRD were because of APKD (Fitzpatrick *et al.*, 1990), while Abbott & Agodoa (2002) carried out a study in the US over 10 years later and estimated the proportion of all those with ESRD because of APKD to be 1.5%.

Garcia Iglesias *et al.* (1983) stated that because of improvements in medical care and diagnostic techniques APKD has been diagnosed earlier and more frequently in recent decades and this, along with therapeutic advances, explains the improvement in kidney and patient survival.

### A.1.1 Genetic Testing

APKD is a single-gene autosomal dominant disorder, passed on from one or other parent through a faulty APKD gene. At present two genes have been identified as causing APKD to develop, PKD1 and PKD2. Reeders *et al.* (1985) showed that the PKD1 locus was closely linked to the  $\alpha$ -globin locus on chromosome 16 and it was later fully sequenced and isolated to location 16p13.3-13.1 by the International Polycystic Kidney Disease Consortium (1995) and Hughes (1995).

Romeo *et al.* (1988) reported a family with APKD but with an absence of linkage between disease mutation and  $\alpha$ -globin, indicating that there is a second mutation. Wright *et al.* (1993) found most APKD families to be due to mutations at the PKD1 locus, but not all. Wright *et al.* (1993) estimated the likelihood that there are 2 APKD loci to 2,514.9 times the likelihood that there is a single locus for APKD. Peters *et al.* (1993) linked the second APKD gene, PKD2, to chromosome 4. This was later isolated to 4q21-23 (Mochizuki *et al.*, 1996).

Torra *et al.* (1996) found that PKD1 mutations account for approximately 85–90% of all people with APKD, while PKD2 accounts for the majority of those that do not carry PKD1. They also noted that there is a possibility of a third gene, PKD3, which has yet to be identified, as Daoust *et al.* (1995) agreed.

The identification of the PKD1 and PKD2 genes enables presymptomatic genetic testing. Linkage analysis is available but is not always suitable as it relies on a large number of affected family members to determine which of the mutations the family carries. If there are enough affected relatives to identify the mutation linkage analysis is very accurate (Harris & Torres, 2006). Sequence analysis is also used to identify PKD1 and PKD2 mutations and the detection rate for this test is found to be about 85% (Rosetti *et al.*, 2001 & 2002).

Sujansky *et al.* (1990) investigated the attitudes of those with APKD and those at risk of APKD to genetic testing. They found that both groups had a good knowledge of APKD and that both groups were concerned about health problems and passing the gene on to children. 97% of the at-risk group said that they would use genetic linkage in order to define their own gene status.

### **A.1.2 Ultrasonographic Scan (USS)**

APKD can also be diagnosed without the use of genetic testing. An ultrasonographic scan (USS) is a test that uses ultrasound to examine the kidneys and determine whether cysts are present and if so the extent of them. Using USS a person is diagnosed as having APKD if there are cysts in both kidneys and one of the kidneys has two or more of these cysts present.

In people with APKD the growth of cysts is age related, therefore the reliability of USS should be age related also. In fact Parfrey *et al.* (1990) and Bear *et al.* (1992) found that the sensitivity of USS is very close to 100% after the age of 30. Demetriou *et al.* (2000) said that although USS was not recommended as a routine diagnostic procedure for those under 14 for people with PKD2, it was 100% reliable in excluding PKD2-related APKD in family members at 50% risk over the age of 30.

Dobin *et al.* (1993) found that the age at diagnosis follows a Normal distribution with mean 20 and standard deviation 15.94. They also found that the penetrance of cysts was over 70% by age 30, over 95% by age 50 and 99% by age 55.

Bear *et al.* (1984) estimated the probability of clinical diagnosis of APKD for those at 50% risk as 0.011, 0.041, 0.115, 0.299 and 0.404 at age 10, 20, 30, 40 and 50 respectively. The probability of ultrasonographic detection of asymptomatic APKD is estimated as 0.222, 0.657 and 0.855 at age 5, 15 and 25, respectively, while the probability of having APKD following a normal ultrasonogram is 0.46, 0.28 and 0.14, respectively, for people at 50% risk, in their 1st, 2nd or 3rd decade.

### **Age at Diagnosis Through Ultrasound**

USS screening is recommended to those at 50% risk of having APKD as it is so

reliable in detecting the cysts produced in people with APKD. One study found that those diagnosed through USS due to family screening had a mean age at diagnosis of 23 years while those diagnosed due to symptoms had a mean age of 40 at diagnosis (Davies *et al.*, 1991).

Torra *et al.* (1996) found that the age of diagnosis using USS was lower for those with PKD1 than those with PKD2, 27.4 years versus 41.4 years, while Wright *et al.* (1993) estimated the age at diagnosis for those with PKD1-related APKD was 25 years compared with 37 years for non-PKD1 related APKD. Other studies found the mean age at diagnosis through USS to lie between 31 and 45.9 years (Sujansky *et al.*, 1990, Gonzalo *et al.*, 1990 and Romão *et al.*, 2006).

Ravine *et al.* (1992) stated that non-PKD1 linked APKD had a much milder phenotype than that of APKD linked to PKD1. Iglesias *et al.* (1997) agreed saying that there is a milder APKD phenotype for those with PKD2 rather than PKD1.

### **A.1.3 Age at Onset of ESRD**

Most of the epidemiological studies that have been carried out have concentrated on the age at onset of ESRD. Table A.1 shows some of the estimates obtained from these studies. Some of the studies in the table have estimates for PKD1 and non-PKD1; these studies were carried out before PKD2 was located and the non-PKD1 group is likely to be the same as the PKD2 group in other studies. This research indicates that the mean age at onset of ESRD is earlier for those with PKD1 rather than PKD2.

Churchill *et al.* (1984) found that the probability of a person with APKD being alive without ESRD was 77% by age 50, 57% by age 58 and 52% by age 73. Gonzalo *et al.* (1990) estimated this probability to be 74% by age 50, 51% by age 58 and 37% by age 70, and Gabow *et al.* (1992) estimated it to be 71% by age 50, 53% by age 58 and 23% by age 70.

Milutinovic *et al.* (1984) found ESRD present in 5% of APKD patients younger than 40 years old, in 33% of patients aged 40 – 49 and in 47% of patients aged 50 years or older.

Table A.1: Mean age at onset of ESRD for people with APKD.

Reference	PKD1	non-PKD1	PKD2	APKD
Parfrey <i>et al.</i> (1990)	56.7	69.4		
Bear <i>et al.</i> (1992)	56.3	68.7		
Torra <i>et al.</i> (1996)	53.5		72.7	
Hateboer <i>et al.</i> (1999)	54.3		74	
Romão <i>et al.</i> (2006)				45.4

Roscoe *et al.* (1993) estimated the mean age at onset of ESRD to be 54.4 for those with APKD and 50.4 for those without APKD, while Abbott & Agodoa (2002) estimated these ages to be 55.24 and 60.56, respectively. It is unclear whether persons with APKD suffer onset of ESRD earlier or later than those without.

Wright *et al.* (1993) estimated that 75% of non-PKD1 related APKD patients had not had ESRD by age 54 compared with 35% of those with PKD1 related APKD.

Gabow *et al.* (1984) found that a number of factors related, independently, with worse renal function at a given age; namely PKD1 gene, younger age at diagnosis, male gender and hypertension.

#### A.1.4 Survival

The survival of those with APKD depends on what form of treatment they undergo, transplant or dialysis, and the form of treatment they undergo depends on their health at the time of ESRD. It was found that, among those that suffer ESRD, healthier people are put on the transplant waiting list (Wolfe *et al.*, 1999). Whether on the waiting list or not people who have ESRD usually underwent dialysis. It was found that the long-term survival was better for those on the waiting list that eventually undergo transplant than for those that did not have a transplant.

Rabbat *et al.* (2000) found that factors associated with the death of patients on the waiting list and the death of patients that underwent transplant were older age and diabetes. Longer time on the waiting list before transplant was also associated

with those that died after transplantation.

Port *et al.* (1993) found that the risk of dying increased early after transplant but, given survival to 365 days, then decreased to a beneficial long-term effect of 60% risk reduction compared with similar patients on dialysis. Meier-Kriesche *et al.* (2001) found that the risk of death decreased by 30% for those that underwent transplantation and by 23% for those that were on the waiting list for transplantation compared with those on dialysis but not on the waiting-list at the time of the study.

McDonald *et al.* (2002) found that cadaveric transplant was associated with an increase in mortality during the first 3 months after the operation, compared with those undergoing dialysis. There was then a reduction in risk of death of 73% at 6 months and 81% at 12 months after the operation.

Murphy *et al.* (2000) studied the differences between hemodialysis and peritoneal dialysis and found that the risk of death was reduced by 17% for those on hemodialysis after 6 months.

Fick *et al.* (1995) found that the main causes of death for those with ESRD were infection (30%) and uremia (28%) before 1975 and cardiac disease, (36%) and infection (24%) after 1975. Roscoe *et al.* (1993) agreed, finding that 34% died from cardiac disease and 20% died from infection, while Harris & Brown (1998) found that 57% died from cardiac disease and 14.3% died from infection.

As with age at onset of ESRD, there have been mixed results from investigations into the differences in survival between those with APKD and those without APKD. Perrone *et al.* (2001) found that those with APKD had a significantly better survival than non-APKD patients, while, Fitzpatrick *et al.* (1990) found that the survival was similar for both APKD and non-APKD groups.

Ravine *et al.* (1992) found the survival to be better for non-PKD1 related APKD patients than for PKD1-related APKD patients, the median survival times being 71.5 and 56.0 years, respectively. Johnson & Gabow (1997) found that those with PKD2 survived longer than those with PKD1 (68 versus 53 years), women had better survival than men (56 versus 52 years), people diagnosed before 30 had worse survival (49 versus 59 years) and those with hypertension before age 35 had

worse survival (51 versus 65 years).

### A.1.5 Gender Difference

Gretz *et al.* (1989) found the median age at ESRD was 52.5 years for men and 58.0 years for women, while Ishikawa *et al.* (2000) found that men with APKD started hemodialysis earlier than women, at age 55.9 compared with age 57.2.

Magistrini *et al.* (2003) estimated the mean age at onset of ESRD for females to be 76 years and for males to be 68.1 years and the median age at CRF (Chronic Renal Failure) to be 72.5 for females and 63.7 for males. Here ESRD is defined as severe chronic renal failure while CRF is defined as moderate chronic renal failure.

Hateboer *et al.* (1999) found that there was a gender difference for those with PKD2 but not for those with PKD1. The median survival with PKD2 was estimated as 71.0 for women and 67.3 for men.

These all suggest that there is some evidence of a gender difference for APKD, with men suffering onset of ESRD earlier than women.

### A.1.6 Hypertension

It has been found that the most important treatable factor associated with faster progression to ESRD in APKD patients is hypertension; the average age of entry into ESRD is 12 years earlier for hypertensive patients than normotensive APKD patients (Gabow *et al.*, 1992). Gonzalo *et al.* (1996) found that prior to renal insufficiency a high arterial pressure makes a significant contribution to renal function deterioration.

## A.2 Critical Illness Insurance Studies

**Gutiérrez & Macdonald (2003)** This study introduced a model for critical illness insurance which has, subsequently, been used in a number of other studies. Figure A.1 shows the CI insurance model, where the CI event for APKD is end-stage renal



disease (ESRD).

The population is divided into three sub-groups:

- Those with no family history, who are not at risk.
- Those at risk because of a family history but who do not carry a mutation.
- Those at risk because of a family history and who do carry a mutation.

The transition intensities in Figure A.1 have to be estimated for each of these subgroups.

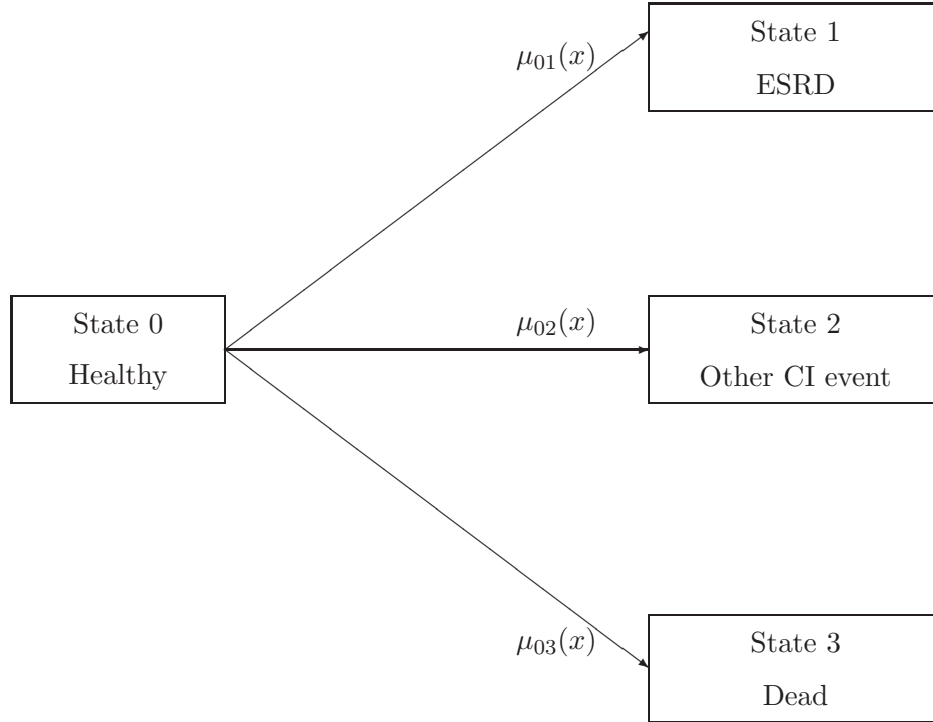


Figure A.1: A multiple state model for APKD in Critical Illness insurance.

This study used data from two sources; Churchill *et al.* (1984) and the U.S. Renal Disease System (USRDS) (1999) to estimate the age-dependent rate of onset of ESRD,  $\mu_{01}(x)$ . The Churchill study, unusually, published the numbers of events, censored cases and persons at risk, which enabled a function to be fitted to the intensity  $\mu_{01}(x)$ . The USRDS study published numbers of cases of ESRD that were due to APKD. These values were used to estimate the onset rates and to fit a function to represent the distribution of age at onset. Churchill *et al.* (1984)

adjusted for any possible ascertainment bias by removing the probands from the analysis.

The estimation of the other intensities,  $\mu_{02}(x)$  and  $\mu_{03}(x)$ , is discussed in Section 6.3

Given all of these intensities the authors calculated level net premiums for a level unit sum assured under a CI insurance contract for a number of ages and policy terms.

In accordance with insurance practice, the level net premiums for mutation carriers were given as a percentage of the level net premiums for non-carriers. This study found that carriers would be uninsurable as the premiums charged would be above 500% of the OR premium rate.

This study also extended the CI insurance model to account for moratoria and adverse selection. They found that if a moratorium covers only genetic test results and not family history, the costs of adverse selection (in terms of increased premiums) are very small. If the moratorium is extended to include family history, the premiums were found to increase slightly.

### **A.3 Available Data and Problems**

There were 46 pedigrees available for APKD noting any available information on whether each person underwent ultrasonographic screening (USS), the result of any genetic testing, when they were diagnosed as having APKD and whether either dialysis or transplantation was required.

The critical illness event for APKD is end-stage renal disease (ESRD) where the kidneys fail and can only be treated through dialysis and transplantation. The problem with the APKD pedigrees is that, although a number of people received a transplant or underwent dialysis, it was not usually noted when a person actually suffered ESRD. For this reason we have not carried out an investigation into the age at onset of ESRD.

Further investigation of the medical notes could be carried out to identify the age at

which ESRD was actually reached. Although knowing that a person has undergone treatment, either transplant or dialysis, indicates that a person is likely to have suffered onset of ESRD it is not conclusive. So, although we could investigate the age at onset of ESRD using the age at which either dialysis or transplant as a surrogate for onset of ESRD, we could not be certain that the results would be plausible.

The other aspect of the pedigrees which could present a problem is that the PKD mutation associated with each family, PKD1 or PKD2, is only known, or noted, for a small number of pedigrees.

Although investigating the critical illness insurance premiums for those known to carry an APKD mutation would be interesting it would have no relevance in the discussion of genetic testing and insurance as APKD has been removed from the list of genetic disorders which are of interest to insurers due to the use of USS and the treatments available.

# Appendix B

## Maximum Likelihood Estimation

### B.1 Maximising the Likelihood

The Nelder-Mead downhill simplex method was used to compute the maximum likelihood estimates by minimising the minus log-likelihood.

The method is outlined here, for more details see Nelder and Mead (1965).

$P$  is defined as a  $(n + 1) \times n$  array containing the points of a  $(n + 1)$ -dimensional simplex, where  $n$  is the number of parameters to be estimated.

The first row of  $P$ ,  $P_0$ , is the set of initial parameter estimates, chosen intuitively. The  $i$ th row of  $P$  is calculated as  $P_i = P_0 + \lambda e_i$ ,  $i = 1, ..n$ , where  $e$  is the set of unit vectors and  $\lambda$  is chosen to reflect the probable scale of the parameters, usually  $\lambda = 1.0$ .

If  $y_i$  is the value of the likelihood at point  $P_i$  then we define  $y_h = \max(y_i)$  and  $y_l = \min(y_i)$ . We also define  $\bar{P}$ , a  $1 \times n$  array, as  $\bar{P}_i = \frac{1}{n} \left( \sum_{j=1}^{n+1} P_{ij} \right)$ , for  $j \neq h$  and

$$\Delta = 2 \frac{|y_h - y_l|}{|y_h| + |y_l| + 1.0e - 10}. \quad (\text{B.1})$$

The minimisation then follows the steps below, with the function being the minus log-likelihood:

A new point,  $P^*$ , is calculated;

$$P^* = (1 + \delta) \bar{P} - \delta P_h. \quad (\text{B.2})$$

If the value of the function at this new point,  $y^*$ , is less than  $y_h$  and more than  $y_l$  the highest point of the simplex,  $P_h$ , then becomes  $P^*$ .

If, however,  $y^*$  is less than  $y_l$  a new point is determined;

$$P^{**} = (1 - \omega) \bar{P} + \omega P^*. \quad (\text{B.3})$$

If the value of the function at this new point,  $y^{**}$ , is less than  $y_l$  then the highest point of the simplex becomes this new point  $P_h = P^{**}$ .

If, however,  $y^{**}$  is bigger than  $y_l$  then the highest point of the simplex  $P_h$  becomes  $P^*$ .

If the value of the function at point  $P^*$  is higher than at all other points of the simplex except for at  $P_h$  then  $P_h = P^*$ . A new point is then determined

$$P^{***} = (1 - \eta) \bar{P} + \eta P_h. \quad (\text{B.4})$$

If the value of the function at this new point,  $y^{***}$ , is less than the minimum of  $y_h$  and  $y^*$  then  $P_h$  becomes  $P^{***}$  otherwise new points are calculated  $P_i = (P_i + P_l) / 2$ .

These steps are repeated until  $\Delta$  is less than some specified value. At this point the optimal values for the parameters are those at the lowest point of the simplex,  $P_l$ .

The optimisation used here is based on a function from *Numerical Recipes in C: The Art of Scientific Computing*.

### B.1.1 Example

To explain the simplex method in more detail we will follow an example for a one-parameter function over two-dimensional space.

We begin with the simplex, in this case a triangle, obtained from the points  $P_0 = (0, 0)$ ,  $P_1 = (0, 1)$ , and  $P_2 = (1, 0)$ , (a) in Figure B.1, where  $\lambda = 1$ . Suppose the point which gives the maximum value of the function,  $P_h$ , is  $P_2$ .

Then  $\bar{P}_1 = (P_{01} + P_{11}) / 2 = (0 + 0) / 2 = 0$  and  $\bar{P}_2 = (0 + 1) / 2 = 0.5$ , so  $\bar{P} = (0, 0.5)$ . With  $\delta = 1$  we then get  $P^* = (-1, 1)$ .

If  $y_l < y^* < y_h$   $P^*$  becomes the new  $P_h$  and a new triangle is made consisting of points  $P_0$ ,  $P_1$  and  $P^*$ , (b) in Figure B.1.

If, however,  $y^*$  is less than  $y_l$ , with  $\omega = 2$ ,  $P^{**} = (-2, 1.5)$ . If  $y^{**}$  is less than  $y_l$ ,  $P^{**}$  replaces  $P_2$ , (c) in Figure B.1, otherwise  $P^*$  does, (b) in Figure B.1.

If  $y^*$  is greater than the value of the function at all points except  $P_h$ , then  $P_h$  becomes either  $P^*$  or the old  $P_h$  whichever gives the lowest value of  $y$ .

Then with  $\eta = 0.5$ , if  $y^*$  is less than  $y_h$  then  $P_h = P^*$  and  $P^{***} = (-0.5, 0.75)$  (e) in Figure B.1. Otherwise  $P_h = P_h$  and  $P^{***} = (0.5, 0.25)$ , (d) in Figure B.1.

If  $y^{***}$  is less than the minimum of  $y_h$  and  $y^*$  then  $P_h = P^{***}$ , otherwise new points are calculated, Step 4.4. These steps are repeated until  $\Delta$  reaches some specified value.

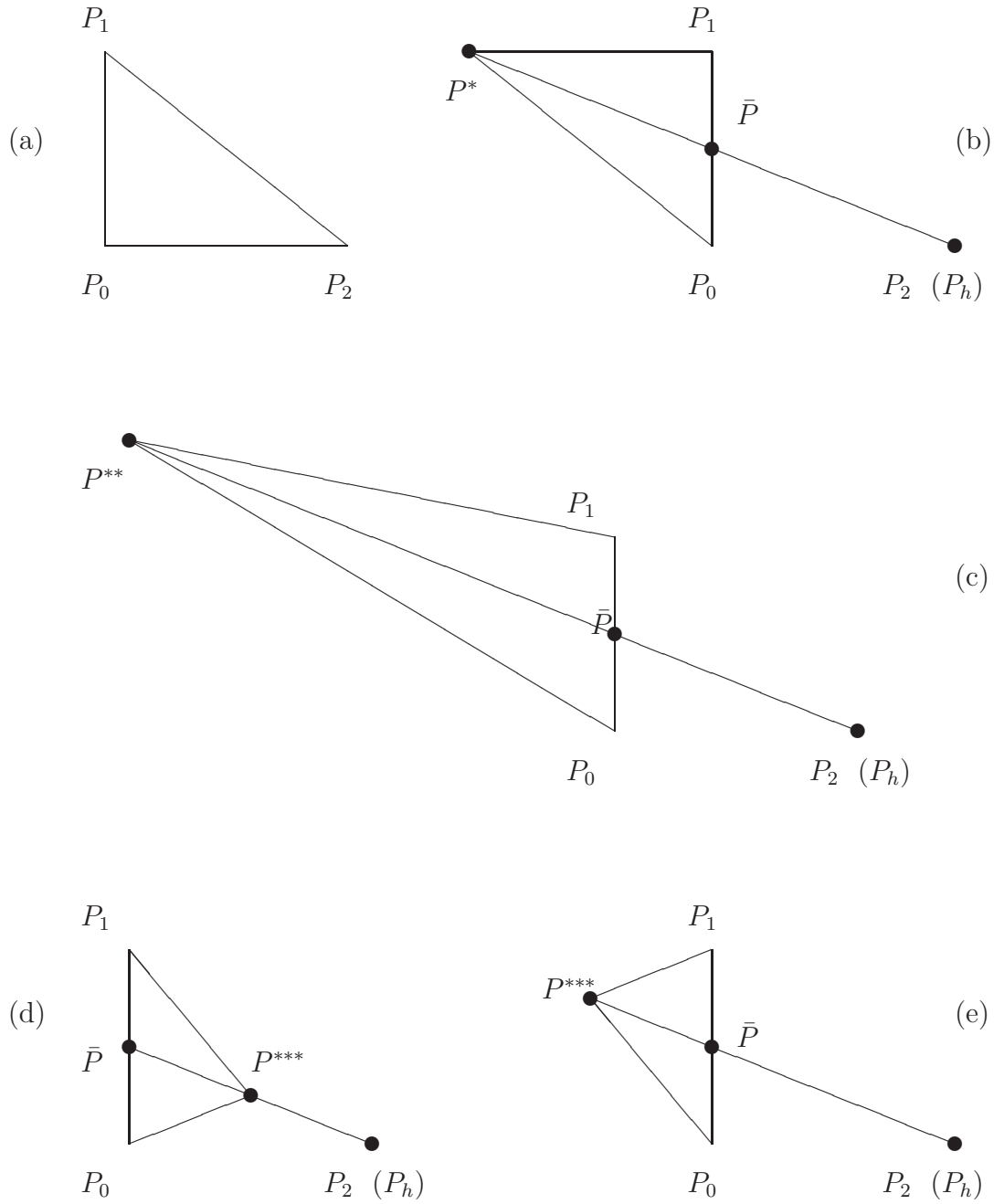


Figure B.1: Example of the Simplex Method for a one parameter function, in two dimensions. (a) is the original simplex. (b) is the 'reflection' step, (c) is the 'expansion' step and (d) and (e) are 'contraction' steps.

# Appendix C

## Other Tools Used

### C.1 Numerical Integration

The main problem occurs when determining the value of the integral. This was overcome by using numerical integration. Two of the methods that could be used are the trapezium method and an acceleration of the Simpson's method.

Trapezium Method:

$$\int_{x_0}^{x_0+\Delta x} f(x)dx = \frac{1}{2} (f(x_0) + f(x_0 + \Delta x)) \Delta x + O(\Delta x^3)$$

Milne's (or Booles) Method:

$$\frac{1}{90} \left[ 7f(x_0) + 32f\left(x_0 + \frac{\Delta x}{4}\right) + 12f\left(x_0 + \frac{\Delta x}{2}\right) + 32f\left(x_0 + \frac{3\Delta x}{4}\right) + 7f(x_0 + \Delta x) \right]$$

Both methods are good approximations to the integral.



## C.2 Polynomial Approximation

### C.2.1 Normal Distribution

There is a polynomial approximation to the Normal available when  $q$  follows a Normal distribution. This is obtained from

$$P(Z < z) = 1 - P(Z > z) \approx 1 - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \sum_{j=1}^5 \frac{b_j}{(1 + pz)^j}$$

Where

$$p = 0.2316419, b_1 = 0.31938153, b_2 = -0.356563782, b_3 = 1.781477937,$$

$$b_4 = -1.821255978, b_5 = 1.330274428$$

This approximation has absolute errors of  $7.5 \times 10^{-8}$ . Giving values very close to the true Standard Normal distribution values. The actual values of the integral were calculated by standardising the time,  $x$ :

$$z = \frac{x - \mu}{\sigma},$$

then using  $z$  to obtain an estimate of integration of the Normal distribution at time  $x$ .

### C.2.2 Gamma Function

Polynomial approximation was also used to determine  $\Gamma(\alpha)$  when  $q$  follows a Gamma distribution as it can be difficult to calculate for values of  $\alpha$  that are not small whole numbers.

$$\Gamma(z) = \left( \frac{\sqrt{2\pi}}{z} \left( p_0 + \sum_{i=1}^6 \frac{p_i}{z+i} \right) \right) (z + 5.5)^{z+0.5} e^{-(z+5.5)}$$

Where

$$p_0 = 1.000000000190015, p_1 = 76.18009172947146, p_2 = -86.50532032941677,$$

$$p_3 = 24.01409824083091, p_4 = -1.231739572450155, p_5 = 0.1208650973866179e - 2,$$

$$p_6 = -0.5395239384953e - 5$$

The function used to calculate  $\Gamma(z)$  in the program is based on a function from *Numerical Recipes in C*.

# References

- ABBOTT, K.C. & AGODOA, L.Y. (2002). Polycystic kidney disease at end-stage renal disease in the United States: Patient characteristics and survival. *Clinical Nephrology*, **57**, 208–214.
- ADAMS, P., FALEK, A. & ARNOLD, J. (1988). Huntington disease in Georgia: age at onset. *American Journal of Human Genetics*, **43**, 695–704.
- AHSAN, H., NEUGUT, A.I., GARBOWSKI, G.C., JACOBSON, J.S., FORDE, K.A., TREAT, M.R. & WAYE, J.D. (1998). Family history of colorectal adenomatous polyps and increased risk for colorectal cancer. *Annals of Internal Medicine*, **128**, 900–905.
- ALARCON, F.J., BURKE, C.A., CHURCH, J.M. & VAN STOLK, R.U. (1999). Familial adenomatous polyposis: Efficacy of endoscopic and surgical treatment for advanced duodenal adenomas. *Diseases of the Colon and Rectum*, **42**, 1533–1536.
- ALONSO, M.E., YESCAS, P., CISNEROS, B., MARTINEZ, C., SILVA, G., OCHOA, A. & MONTAÑEZ, C. (1997). Analysis of the (CAG)<sub>n</sub> repeat causing Huntington's disease in a Mexican population. *Clinical Genetics*, **51**, 225–230.
- ANDERSEN P.K., BORGAN, O, GILL, R.D. & KEIDING, N. (1993). *Statistical Models Based on Counting Processes*. Springer.
- ANDERSON, K.E., EASTON, D.F., MATTHEWS, F.E. & PETO, J. (1992). Cancer mortality in the first degree relatives of young breast cancer patients. *British Journal of Cancer*, **66**, 599–602.
- ANDREW, S., GOLDBERG, P., KREMER, B., TELENUS, H., THIELMANN, J., ADAM, S., STARR, E., SQUITIERI, F., LIN, B., KALCHMAN, M., GRAHAM,

R. & HAYDEN, M. (1993). The relationship between trinucleotide (CAG) repeats length and clinical features of Huntington's disease. *American Journal of Human Genetics*, **43**, 695–704.

ANGLICAN BREAST CANCER STUDY GROUP (2000). Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. *British Journal of Cancer*, **83**, 1301–1308.

ANTONIOU, A.C., GAYTHER, S.A., STRATTON, J.F., PONDER, B.A.J. & EASTON, D.F. (2000). Risk models for familial ovarian and breast cancer. *Genetic Epidemiology*, **18**, 173–190.

ANTONIOU, A.C., PHAROAH, P.D., McMULLAN, G., DAY, N.E., STRATTON, M.R., PETO, J., PONDER, B.J. & EASTON D.F. (2002). A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *British Journal of Cancer*, **86**, 76–83.

ANTONIOU, A.C., PHAROAH, P.D., NAROD, S., RISCH, H.A., EYFJORD, J.E., HOPPER, J.L., LOMAN, N., OLSSON, H., JOHANSSON, O., BORG, A., PASINI, B., RADICE, P., MANOUKIAN, S., ECCLES, D.M., TANG, N., OLAH, E., ANTON-CULVER, H., WARNER, E., LUBINSKI, J., GRONWALD, J., GORSKI, B., TULINIUS, H., THORLACIUS, S., EEROLA, H., NEVANLINNA, H., SYRJAKOSKI, K., KALLIONIEMI, O.P., THOMPSON, D., EVANS, C., PETO, J., LALLOO, F., EVANS, D.G. & EASTON, D.F. (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. *American Journal of Human Genetics*, **72**, 1117–1130.

ANTONIOU, A., PHAROAH, P.D.P., NAROD, S., RISCH, H.A., EYFJORD, J.E., HOPPER, J.L., LOMAN, N., OLSSON, H., JOHANSSON, O., BORG, A., PASINI, B., RADICE, P., MANOUKIAN, S., ECCLES, D.M., TANG, N., OLAH, E., ANTON-CULVER, H., WARNER, E., LUBINSKI, J., GRONWALD, J., GORSKI, B., TULINIUS, H., THORLACIUS, S., EEROLA, H., NEVANLINNA, H., SYRJÄKOSKI, K., KALLIONIEMI, O.P., THOMPSON, D., EVANS, C., PETO, J., LALLOO, F., EVANS, D.G. & EASTON, D.F. (2005). Breast and ovarian cancer risks to carriers of the BRCA1 5382insC and 185delAG and BRCA2 6174delT mutations: A

combined analysis of 22 population based studies. *Journal of Medical Genetics*, **42**, 602–603.

ANTONIOU, A.C., DUROCHER, F., SMITH, P., SIMARD, J., INHERIT BRCA5 PROGRAM MEMBERS & EASTON, D.E. (2006). BRCA1 and BRCA2 mutation predictions using the BOADICEA and BRCAPRO models and penetrance estimation in high-risk French-Canadian families . *Breast Cancer Research*, **8**, R3.

ARMSTRONG, J.G., DAVIES, D.R., GUY, S.P., FRAYLING, I.M. & EVANS, D.G.R. (1997). APC mutations in familial adenomatous polyposis families in the northwest of England. *Human Mutation*, **10**, 376–380.

ARMSTRONG, K., SCHWARTZ, J.S., RANDALL, T., RUBIN, S.C. & WEBER, B. (2004). Hormone replacement therapy and life expectancy after prophylactic oophorectomy in women with BRCA1/2 mutations: A decision analysis. *Journal of Clinical Oncology*, **22**, 1045–1054.

ARVANITIS, M.L., JAGELMAN, D.G., FAZIO, V.W., LAVERY, I.C. & MCGANNON, E. (1990). Mortality of patients with familial adenomatous polyposis. *Diseases of the Colon and Rectum*, **33**, 639–642.

ASSOCIATION OF BRITISH INSURERS (2005). *Concordat and moratorium on genetics and insurance*. ABI, London.

ASSOCIATION OF BRITISH INSURERS (1997). Unpublished Document. .

BARRON, L.H., WARNER, J.P., PORTEOUS, M., HOLLOWAY, S., SIMPSON, S., DAVIDSON, R. & BROCK, D.J.H. (1993). A study of the Huntington’s disease associated trinucleotide repeat in the Scottish population. *Journal of Medical Genetics*, **30**, 1003–1007.

BEAR, J.C., MCMANAMON, P., MORGAN, J., PAYNE, R.H., LEWIS, H., GAULT, M.H. & CHURCHILL, D.N. (1984). Age at clinical onset and at ultrasonographic detection of adult polycystic kidney disease: Data for genetic counselling. *American Journal of Medical Genetics*, **18**, 45–53.

BEAR, J.C., PARFREY, P.S., MORGAN, J.M., MARTIN, C.J. & CRAMER, B.C. (1992). Autosomal dominant polycystic kidney disease: New information for genetic

- counselling. *American Journal of Medical Genetics*, **43**, 548–553.
- BELL, J. (1934). Huntington's chorea. *Treasury of Human Inheritance*, **IV**, 1–67.
- BERNSTEIN, I.T. & BÜLOW, S. (2005). Hereditary colorectal cancer. *Ugeskr Laeger*, **167**, 4159–4163.
- BERTARIO, L., RUSSO, A., SALA, P., EBOLI, M., GIAROLA, M., D'AMICO, F., GISMONDI, V., VARESCO, L., PIEROTTI, M.A. & RADICE, P., ON BEHALF OF THE HEREDITARY COLORECTAL TUMOURS REGISTRY (2001). Genotype and phenotype factors as determinants of desmoid tumors in patients with familial adenomatous polyposis. *International Journal of Cancer*, **95**, 102–107.
- BERTARIO, L., RUSSO, A., SALA, P., VARESCO, L., GIAROLA, M., MONDINI, P., PIEROTTI, M., SPINELLI, P. & RADICE, P., FOR THE HEREDITARY COLORECTAL TUMOR REGISTRY (2003). Multiple approach to the exploration of genotype-phenotype correlations in familial adenomatous polyposis. *Journal of Clinical Oncology*, **21**, 1698–1707.
- BINEDELL, J., SOLDAN, J.R. & HARPER, P.S. (1998). Predictive testing for Huntington's disease: I. Predictors of uptake in South Wales. *Clinical Genetics*, **54**, 477–488.
- BISGAARD, M.L., FENGER, K., BÜLOW, S., NIEBUHR, E. & MOHR, J. (1994). Familial adenomatous polyposis (FAP): Frequency, penetrance and mutation rate. *Human Mutation*, **3**, 121–125.
- BJÖRK, J., AKERBRANT, H., ISELIUS, L., ALM, T. & HULTCRANTZ, R. (1999). Epidemiology of familial adenomatous polyposis in Sweden: changes over time and differences in phenotype between males and females. *Scandinavian Journal of Gastroenterology*, **34**, 1230–1235.
- BODMER, W.F., BAILEY, C.J., BODMER, J., BUSSEY, H.J., ELLIS, A., GORMAN, P., LUCIBELLO, F.C., MURDAY, V.A., RIDER, S.H., SCAMBLER, P., SHEER, D., SOLOMON, E. & SPURR, N.K. (1987). Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature*, **328**, 614–616.
- BOEHNKE, M. & GREENBERG, D.A. (1984). The effects of conditioning on pro-

bands to correct for multiple ascertainment. *American Journal of Human Genetics*, **36**, 1298–1308.

BOLT, J.M. (1970). Huntington's chorea in the West of Scotland. *British Journal of Psychiatry*, **116**, 259–270.

BONELLI, L., MARTINES, S., CONIO, M., BRUZZI, P. & ASTE, H. (1988). Family history of colorectal cancer as a risk factor for benign and malignant tumours of the large bowel. A case-control study. *International Journal of Cancer*, **41**, 513–517.

BRACKENRIDGE, C. (1971). A genetic and statistical study of some sex-related factors in Huntington's disease. *Clinical Genetics*, **2**, 267–286.

BRANDT, J., BYLSMA, F.W., GROSS, R., STINE, O.C., RANE, N. & ROSS, C.A. (1996). Trinucleotide repeat length and clinical progression in Huntington's disease. *Neurology*, **46**, 527–531.

BRINKMAN, R., MEZEI, M., THEILMANN, J., ALMQVIST, E. & HAYDEN, M. (1997). The likelihood of being affected with Huntington disease by a particular age, for a specific CAG size. *American Journal of Human Genetics*, **60**, 1202–1210.

BROTHERS, C.R.D (1964). Huntington's chorea in Victoria and Tasmania. *Journal of Neurological Sciences*, **1**, 405–420.

BÜLOW, S. (1986). Clinical features in familial adenomatous polyposis coli, Results of the Danish polyposis register. *Diseases of the Colon and Rectum*, **29**, 102–107.

BÜLOW, S., ALM, T., FAUSA, O., HULTERANTZ, R., JÄRVINEN, H., VASEN, H. & DAF PROJECT GROUP (1995). Duodenal adenomatosis in familial adenomatous polyposis. *International Journal of Colorectal Disease*, **10**, 43–46.

BÜLOW (1996). The incidence rate of familial adenomatous polyposis. Results from the Danish polyposis register. *International Journal of Colorectal Diseases*, **11**, 88–91.

BÜLOW, C., VASEN, H., JÄRVINEN, H., BJÖRKA, J., BISGAARD, M.L. & BÜLOW, S. (2000). Ileorectal Anastomosis Is Appropriate for a Subset of Patients

- With Familial Adenomatous Polyposis. *Gastroenterology*, **199**, 1454–1460.
- BÜLOW, S. (2003). Results of national registration of familial adenomatous polyposis. *Gut*, **52**, 742–746.
- BÜLOW, S., BJÖRK, J., CHRISTENSEN, I.J., FAUSA, O., JÄRVINEN, H., MOESGAARD, F., VASEN, H.F.A., THE DAF STUDY GROUP (2004). Duodenal adenomatosis in familial adenomatous polyposis. *Gut*, **53**, 381–386.
- BURN, J., CHAPMAN, P., DELHANTY, J., WOOD, C., LALLOO, F., CACHON-GONZALEZ, M.B., TSIOPRA, K., CHURCH, W., RHODES, M. & GUNN, A. (1991). The UK Northern Region genetic register for familial adenomatous polyposis coli: use of age of onset, congenital hypertrophy of the retinal pigment epithelium and DNA markers in risk calculations. *Journal of Medical Genetics*, **28**, 289–296.
- BURT, R.W. (2000). Colon cancer screening. *Gastroenterology*, **119**, 837–853.
- BUSSEY, H.J.R. (1975). *Familial polyposis coli. Family studies, histopathology, differential diagnosis and results of treatment*. The John Hopkins University Press, Baltimore.
- CANNINGS, C. & THOMPSON, E.A. (1977). Ascertainment in the sequential sampling of pedigrees. *Clinical Genetics*, **12**, 208–212.
- CHUNG, D.C. (2000). The genetic basis of colorectal cancer: Insights into critical pathways of tumorigenesis. *Gastroenterology*, **119**, 854–865.
- CHURCHILL, D.N., BEAR, J.C., MORGAN, J., PAYNE, R.H., MCMANAMON, P.J. & GAULT, M.H. (1984). Prognosis of adult onset polycystic kidney disease re-evaluated. *Kidney International*, **26**, 190–193.
- CLARK, S.K. & PHILLIPS, R.K.S. (1996). Desmoids in familial adenomatous polyposis. *British Journal of Surgery*, **83**, 1494–1504.
- CLAUS, E.B., RISCH, N. & THOMPSON, W.D. (1991). Genetic analysis of breast cancer in the cancer and steroid hormone study. *American Journal of Human Genetics*, **48**, 232–242.
- CLAUS, E.B., RISCH, N. & THOMPSON, W.D. (1993). The calculation of breast



cancer risk for women with a first degree family history of ovarian cancer. *Breast Cancer Research and Treatment*, **28**, 115–120.

CLAUS, E.B., RISCH, N. & THOMPSON, W.D. (1994). Autosomal dominant inheritance of early-onset breast cancer. *Cancer*, **73**, 643–651.

CLAUS, E.B., SCHILDKRAUT, J., IVERSEN, E.S., BERRY, D. & PARMIGIANI, G. (1998). Effect of BRCA1 and BRCA2 on the association between breast cancer risk and family history. *Journal of the National Cancer Institute*, **90**, 1824–1829.

COLEMAN, M., BABB, P., DAMIECKI, P., GROSCLAUDE, P.C., HONJO, S., JONES, J., KNERER, G., PITARD, A., QUINN, M.J., SLOGGETT, A. & DE STAVOLA, B.L. (1999). *Cancer survival trends for England and Wales, 1971-1995: Deprivation and NHS Region*. Office of National Statistics, London.

COLLABORATIVE GROUP ON HORMONAL FACTORS IN BREAST CANCER (WRITING COMMITTEE: BERAL, V., BULL, D., DOLL, R. & REEVES, G.) (2001). Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without disease. *The Lancet*, **358**, 1389–1399.

COUCH, F.J., DESHANO, M.L., BLACKWOOD, M.A., CALZONE, K., STOPFER, J., CAMPEAU, L., GANGULY, A., REBBECK, T. & WEBER, B.L. (1997). BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer. *The New England Journal of Medicine*, **336**, 1409–1415.

CRAUFURD, D., DODGE, A., KERZIN-STORRAR, L. & HARRIS, R. (1989). Uptake of presymptomatic predictive testing for Huntington's disease. *Lancet*, **ii**, 603–605.

CREIGHTON, S., ALMQVIST, E.W., MACGREGOR, D., FERNANDEZ, B., HOGG, H., BEIS, J., WELCH, J.P., RIDDELL, C., LOKKESMOE, R., KHALIFA, M., MACKENZIE, J., SAJOO, A., FARRELL, S., ROBERT, F., SHUGAR, A., SUMMERS, A., MESCHINO, W., ALLINGHAM-HAWKINS, D., CHIU, T., HUNTER, A., ALLANSON, J., HARE, H., SCHWEEN, J., COLLINS, L., SANDERS, S., GREENBERG, C., CARDWELL, S., LEMIRE, E., MACLEOD, P. & HAYDEN, M.R. (2003). Predictive, pre-natal and diagnostic genetic testing for Huntington's disease: The

experience in Canada from 1987 to 2000. *Clinical Genetics*, **63**, 462–475.

CRONER, R.S., BRUECKL, W.M., REINGRUBER, B., HOHENBERGER, W. & GUENTHER, K. (2005). Age and manifestation related symptoms in familial adenomatous polyposis. *BMC Cancer*, **5**, 24.

DAOUST, M.C., REYNOLDS, D.M., BICHET, D.G. & SOMLO, S. (1995). Evidence for a third genetic locus for autosomal dominant polycystic kidney disease. *Genomics*, **25**, 733–736.

DAVIES, F., COLES, G.A., HARPER, P.S., WILLIAMS, A.J., EVANS, C. & COCHLIN, D. (1991). Polycystic kidney disease re-evaluated: A population based study. *Quarterly Journal of Medicine, New Series* **79**, 477–485.

DAYKIN, C.D., AKERS, D.A., MACDONALD, A.S., MCGLEENAN, T., PAUL, D. & TURVEY, P.J. (2003). Genetics and Insurance - Some Social Policy Issues (with discussions). *British Actuarial Journal*, **9**, 787–874.

DEMETRIOU, K., TZIAKOURI, C., ANNINO, K., ELEFThERIOU, A., KOPTIDES, M., NICOLAOU, A., DELTAS, C.C. & PIERIDES, A. (2000). Autosomal dominant polycystic kidney disease - type 2. Ultrasound, genetic and clinical correlations. *Nephrology Dialysis Transplantation*, **15**, 205–211.

DEWHURST, K., OLIVER, J.E. & MCKNIGHT, L. (1970). Huntington's disease. *British Journal of Psychiatry*, **116**, 255–258.

DINANI, A., GRIMSHAW, D., ROBJOHNS, N., SOMERVILLE, S., SPRY, A. & STAFFURTH, J. (2000). *A critical review: Report of the critical illness healthcare study group*. London: Staple Inn Actuarial Society.

DOBIN, A., KIMBERLING, W.J., PETTINGER, W., BAILEY-WILSON, J.E., SHUGART, Y.Y. & GABOW, P. (1993). Segregation analysis of autosomal dominant polycystic kidney disease. *Genetic Epidemiology*, **10**, 189–200.

EASTON, D.F., FORD, D., BISHOP, T. & THE BREAST CANCER LINKAGE CONSORTIUM (1995). Breast and ovarian cancer incidence in BRCA1-mutation carriers. *American Journal of Human Genetics*, **56**, 265–271.

- ELSTON, R.C. (1973). Ascertainment and age of onset in pedigree analysis. *Human Heredity*, **23**, 105–112.
- EVANS, D.G.R., GUY, S.P., THAKKER, N., ARMSTRONG, J.G., DODD, C., DAVIES, D.R., BABBS, C., CLANCY, T., WARNES, T., SLOAN, P., TAYLOR, T.V. & HARRIS, R. (1993). Non-penetrance and late appearance of polyps in families with familial adenomatous polyposis. *Gut*, **34**, 1389–1393.
- EWENS, W.J. & SHUTE, N.C.E. (1986). A resolution of the ascertainment sampling problem. I. Theory. *Theoretical Population Biology*, **30**, 388–412.
- FARRER, L. & CONNEALLY, P. (1985). A genetic model for age at onset in Huntington disease. *American Journal of Human Genetics*, **37**, 350–357.
- FEIGIN, A., KIEBURTZ, K., BORDWELL, K., COMO, P., STEINBERG, K., SOTACK, J., ZIMMERMAN, C., HICKEY, C., ORMEN, C. & SHOULSON, I. (1995). Functional decline in Huntington's disease. *Movement Disorders*, **10**, 211–214.
- FICK, G.M., JOHNSON, A.M., HAMMOND, W.S. & GABOW, P.S. (1995). Causes of death in autosomal dominant polycystic kidney disease. *Journal of the American Society of Nephrology*, **5**, 2048–2056.
- FITZPATRICK, P.M., TORRES, V.E., CHARBONEAU, J.W., OFFORD, K.P., HOLLEY, K.E. & ZINCKE, H. (1990). Long-term outcome of renal transplantation in autosomal dominant polycystic kidney disease. *American Journal of Kidney Diseases*, **XV**, 535–543.
- FOLSTEIN, S.E., CHASE, G.A., WAHL, W.E., MCDONNELL, A.M. & FOLSTEIN, M.F. (1987). Huntington disease in Maryland: Clinical aspects of racial variation. *American Journal of Human Genetics*, **41**, 168–179.
- FORD, D., EASTON, D.F., STRATTON, M., NAROD, S., GOLDGAR, D., DEVILLE, P., BISHOP, D.T., WEBER, B., LENOIR, G., CHANG-CLAUDE, J., SOBOLEW, H., TEARE, M.D., STRUEWING, J., ARASON, A., SCHERNECK, S., PETO, J., REBBECK, T.R., TONIN, P., NEUHAUSEN, S., BARKARDOTTIR, R., EYFJORD, J., LYNCH, H., PONDER, B.A.J., GAYTHER, S.A., BIRCH, J.M., LINDBLOM, A., STOPPA-LYONNET, D., BIGNON, Y., BORG, A., HAMANN, U., HAITEs, N.,

- SCOTT, R.J., MAUGARD, C.M., VASEN, H., SEITZ, S., CANNON-ALBRIGHT, L.A., SCHOFIELD, A., ZELADA-HEDMAN, M. & THE BREAST CANCER LINKAGE CONSORTIUM (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *American Journal of Human Genetics*, **62**, 676–689.
- FOROUD, T., GRAY, J., IVASHINA, J. & CONNEALLY, M. (1999). Differences in duration of Huntington's disease based on age at onset. *Journal of Neurology, Neurosurgery and Psychiatry*, **66**, 52–56.
- FRANK, T.S., MANLEY, S.A., OLOPADE, O.I., CUMMINGS, S., GARBER, J.E., BERNHARDT, B., ANTMAN, K., RUSSO, D., WOOD, M.E., MILLINEAU, L., ISAACS, C., PESHKIN, B., BUYS, S., VENNE, V., ROWLEY, P.T., LOADER, S., OFFIT, K., ROBSON, M., HAMPEL, H., BRENER, D., WINER, E.P., CLARK, S., WEBER, B., STRONG, L.C., THOMAS, A. (1998). Sequence analysis of BRCA1 and BRCA2: Correlation of mutations with family history and ovarian cancer risk. *Journal of Clinical Oncology*, **16**, 2417–2425.
- FUCHS, C.S., GIOVANNUCCI, E.L., COLDITZ, G.A., HUNTER, D.J., SPEIZER, F.E. & WILLETT, W.C. (1994). A prospective study of family history and risk of colorectal cancer. *New England Journal of Medicine*, **331**, 1669–1674.
- FURTADO, S., SUCHOWERSKY, O., REWCASTLE, B., GRAHAM, L., KLIMEK, M.L. & GARBER, A. (1996). Relationship between trinucleotide repeats and neuropathological changes in Huntington's disease. *Annals of Neurology*, **39**, 132–136.
- GABOW, P.A., IKLÉ, D.W. & HOLMES, J.H. (1984). Polycystic kidney disease: Prospective analysis of nonazotemic patients and family members. *Annals of Internal Medicine*, **101**, 238–247.
- GABOW, P.A., JOHNSON, A.M., KAEHNY, W.D., KIMBERLING, W.J., LEZOTTE, D.C., DULEY, I.T. & JONES, R.H. (1992). Factors affecting the progression of renal disease in autosomal dominant polycystic kidney disease. *Kidney International*, **41**, 1311–1319.
- GARCIA IGLESIAS, C., TORRES, V.E., OFFORD, K.P., HOLLEY, K.E., BEARD, C.M. & KURKLAND, L.T. (1983). Epidemiology of adult polycystic kidney dis-

ease, Olmsted County, Minnesota: 1935 - 1980. *American Journal of Kidney Diseases*, **II**, 630–639.

GENETICS AND INSURANCE COMMITTEE (2000). *First Annual Report*. Department of Health, London.

GONZALO, A., RIVERA, M., QUEREDA, C. & ORTUÑO, J. (1990). Clinical features and prognosis of adult polycystic kidney disease. *American Journal of Nephrology*, **10**, 470–474.

GONZALO, A., GALLEGU, A., RIVERA, M., ORTE, L. & ORTUÑO (1996). Influence of hypertension on early renal insufficiency in autosomal dominant polycystic kidney disease. *Nephron*, **72**, 225–230.

GRETZ, N., ZEIER, M., GEBERTH, S., STRAUCH, M. & RITZ, E. (1989). Is gender a determinant for evolution of renal failure? A study in autosomal dominant polycystic kidney disease. *American Journal of Kidney Diseases*, **XIV**, 178–183.

GRODEN, J., THLIVERIS, A., SAMOWITZ, W., CARLSON, M., GELBERT, L., ALBERTSEN, H., JOSLYN, G., STEVENS, J., SPIRIO, L., ROBERTSON, M., SARGEANT, L., KRAPCHO, K., WOLFF, E., BURT, R., HUGHES, J.P., WARRINGTON, J., MCPHERSON, J., WASMUTH, J., LEPASLIER, D., ABDERRAHIM, H., COHEN, D., LEPPERT, M. & WHITE, R. (1991). Identification and characterisation of the familial adenomatous polyposis coli gene. *Cell*, **66**, 589–600.

GUI, E.H. & MACDONALD, A.S. (2002). A Nelson-Aalen Estimate of of the Incidence Rates of Early-Onset Alzheimer's Disease Associated With the Presenilin-1 Gene. *ASTIN Bulletin*, **32**, 1–42.

GUI, E.H., LU, B., MACDONALD, A.S., WATERS, H.R & WEKWETE, C.T. (2006). The Genetics and Breast and Ovarian Cancer III: A New Model of Family History With Insurance Applications. *06/1*.

GUSELLA, J.F., WEXLER, N.S., CONNEALLY, P.M., NAYLOR, S.L., ANDERSON, M.A, TANZI, R.E., WATKINS, P.C., OTTINA, K., WALLACE, M.R., SAKAGUCHI, A.Y., YOUNG, A.B., SHOULSTON, I., BONILLA, E. & MARTIN, J.B. (1983). A polymorphic DNA marker genetically linked to Huntington's Disease.

*Nature*, **306**, 234–238.

GUTIÉRREZ, M.C. & MACDONALD, A.S. (2002a). Huntington's Disease and Insurance I: A Model of Huntington's Disease. *GIRC Research Report 02/3*.

GUTIÉRREZ, M.C. & MACDONALD, A.S. (2002b). Huntington's Disease and Insurance I: Critical Illness Insurance and Life Insurance. *GIRC Research Report 02/4*.

GUTIÉRREZ, M.C. & MACDONALD, A.S. (2003). Adult polycystic kidney disease and critical illness insurance. *North American Actuarial Journal*, **7:2**, 93–115.

GUTIÉRREZ, M.C. & MACDONALD, A.S. (2004). Huntington's disease, critical illness insurance and life insurance. *Scandinavian Actuarial Journal*, **2004**, 279–311.

HALL, J.M., LEE, M.K., NEWMAN, B., MORROW, J.E., ANDERSON, L.A., HUEY, B. & KING, M-C. (1990). Linkage of early onset familial breast cancer to chromosome 17q21. *Science*, **250**, 1684–1689.

HARPER, P.S., LIM, C. & CRAUFURD, D. (2000). Ten years of presymptomatic testing for Huntington's disease: the experience of the UK Huntington's Disease Prediction Consortium.. *Journal of Medical Genetics*, **37**, 567–71.

HARPER, P.S. (1996). *Huntington's disease*. W.B. Saunders.

HARRIS, S.A.C. & BROWN, E.A. (1998). Patients surviving more than 10 years on hemodialysis. The natural history of the complications of treatment. *Nephrology Dialysis Transplantation*, **13**, 1226–1233.

HARRIS, P.C. & TORRES, V.E. (Updated June 2006). Autosomal Dominant Polycystic Kidney Disease. In *GeneReviews at GeneTests: Medical Genetics Information Resource (database online)*. Copyright: University of Washington, Seattle. (1997-2006). Available at <http://www.genetests.org>.

HATEBOER, N., v DIJK, M.A., BOGDANOVA, N., COTO, E., SAGGAR-MALIK, A.K., SAN MILLAN, J.L., TORRA, R., BREUNING, M., RAVINE, D., THE EUROPEAN PKD1-PKD2 STUDY GROUP (1999). Comparison of phenotypes of poly-

cystic kidney disease types 1 and 2. *Lancet*, **353**, 103–107.

HEISKANEN, I., LUOSTARINEN, T. & JÄRVINEN, H.J. (2000). Impact of screening examinations on survival in familial adenomatous polyposis. *Scandinavian Journal of Gastroenterology*, **35**, 1284–1287.

HODGE, S.E. & VIELAND, V.J. (1996). The essence of single ascertainment. *Genetics*, **144**, 1215–1223.

HOGERVORST, F.B., *et al.* (1995). Rapid detection of BRCA1 mutations by the protein truncation test. *Nature Genetics*, **10**, 208–212.

HUGHES, J., WARD, C.J., PERAL, B., ASPINWALL, R., CLARK, K., SAN MILLAN, J.L., GAMBLE, V. & HARRIS, P.C. (1995). The polycystic kidney disease 1 (PKD1) gene encodes a novel protein with multiple cell recognition domains. *Nature Genetics*, **10**, 151–160.

THE HUNTINGTON'S DISEASE COLLABORATIVE RESEARCH GROUP (1993). A novel gene containing a trinucleotide repeat that is extended and unstable on Huntington's Disease chromosomes. *Cell*, **72**, 971–983.

IGLESIAS, D.M., MARTIN, R.S., FRAGA, A., VIRGINILLO, M., KORNBLIHTT, A.R., ARRIZURIETA, E., VIRIBAY, M., SAN MILLÁN, J.L., HERRERA, M. & BERNATH, V. (1997). Genetic heterogeneity of autosomal dominant polycystic kidney disease in Argentina. *Journal of Medical Genetics*, **34**, 827–830.

ILLARIOSHKIN, S.N., IGARASHI, S., ONODERA, O., MARKOVA, E.D., NIKOLSKAYA, N.N., TANAKA, H., CHABRASHWILI, T.Z., INSAROVA, N.G., ENDO, K., IVANOVA-SMOLENSKAYA, I.A. & TSUJI, S. (1994). Trinucleotide repeat length and rate of progression of Huntington's disease. *Annals of Neurology*, **36**, 630–635.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2004). Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931–945.

INTERNATIONAL POLYCYSTIC KIDNEY DISEASE CONSORTIUM (1995). Polycystic kidney disease: The complete structure of the PKD1 gene and its protein. *Cell*, **81**, 289–298.



- ISHIKAWA, I., MAEDA, K., NAKAI, S. & KAWAGUCHI, Y. (2000). Gender difference in the mean age at the induction of hemodialysis in patients with autosomal dominant polycystic kidney disease. *American Journal of Kidney Diseases*, **35**, 1072–1075.
- JAKAB, K., GÁRDIÁN, G., ENDREFFY, E., KALMÁR, T., BACHRATI, C., VÉCSEI, L. & RASKÓ, I. (1999). Analysis of CAG repeat expansion in Huntington's disease gene (IT 15) in a Hungarian population. *European Neurology*, **41**, 107–110.
- JÄRVINEN, H.J. (1985). Time and type of prophylactic surgery for familial adenomatous coli. *Annals of Surgery*, **202**, 93–97.
- JÄRVINEN, H.J. (1992). Epidemiology of familial adenomatous polyposis in Finland: impact of family screening on the colorectal cancer rate and survival. *Gut*, **33**, 357–360.
- JÄRVINEN, H.J. (2003). Genetic testing for polyposis: practical and ethical aspects. *Gut*, **52**(Supplement II), ii9–ii22.
- JÄRVINEN, H.J. (2004). Hereditary cancer: Guidelines in clinical practice. Colorectal cancer genetics. *Annals of Oncology*, **15** (Supplement 4), 127–131.
- JOHNSON, A.M. & GABOW, P.A. (1997). Identification of patients with autosomal dominant polycystic kidney disease at highest risk for end-stage renal disease. *Journal of the American Society of Nephrology*, **8**, 1560–1567.
- JONES, M.B. & PHILLIPS, C.R. (1970). Affected parent and age at onset of Huntington's chorea. *Journal of Medical Genetics*, **7**, 20–21.
- KEHOE, P., KRAWCZAK, M., HARPER, P.S., OWEN, M.J. & JONES, A.L. (1999). Age of onset of Huntington's disease: sex specific influence of apolipoprotein E genotype and normal CAG repeat length. *Journal of Medical Genetics*, **36**, 108–111.
- KIEBURTZ, K., MACDONALD, M., SHIH, C., FEIGIN, A., STEINBERG, K., BORDWELL, K., ZIMMERMAN, C., SRINIDHI, J., SOTACK, J., GUSELLA, J. & SHOULSON, I. (1994). Trinucleotide repeat length and progression of illness in



Huntington's disease. *Journal of Medical Genetics*, **31**, 872–874.

KINZLER, K.W., NILBERT, M.C., SU, L.K., VOGELSTEIN, B., BRYAN, T.M., LEVY, D.B., SMITH, K.J., PREISINGER, A.C., HEDGE, P., McKECHNIE, D., FINNIEAR, R., MARKHAM, A., GROFFEN, J., BOGUSKI, M.S., ALTSCHUL, S.F., HORII, A., ANDO, H., MIYOSHI, Y., MIKI, Y., NISHISHO, I. & NAKAMURA, Y. (1991). Identification of FAP locus genes from chromosome 5q21. *Science*, **253**, 661–665.

KRAWCZAK, M., BOCKEL, B., SANDKUIJL, L., THIES, U., FENTON & HARPER, P.S. (1991). Covariate-dependent age-at-onset distributions for Huntington disease. *American Journal of Human Genetics*, **49**, 735–745.

LALLOO, F., VARLEY, J., MORAN, A., ELLIS, D., O'DAIR, L., PHAROAH, P., ANTONIOU, A., HARTLEY, R., SHENTON, A., SEAL, S., BULMAN, B., HOWELL, A., GARETH, D. & EVANS, R. (2006). BRCA1, BRCA2 and TP53 mutations in very early-onset breast cancer with associated risks to relatives. *European Journal of Cancer*, **42**, 1143–1150.

LANGBEHN, D.R., FALUSH, R.R., PAULSEN, D. & HAYDEN, M.R. ON BEHALF OF AN INTERNATIONAL HUNTINGTON'S DISEASE COLLABORATIVE GROUP (2004). A new model for prediction of the age of onset and penetrance for Huntington's Disease based on CAG length. *Clinical Genetics*, **65**, 267–277.

LEIGH, T.S. (1990). Underwriting – a dying art? (with discussion). *Journal of the Institute of Actuaries*, **117**, 443–531.

LEMAIRE, J., SUBRAMANIAN, K., ARMSTRONG, K. & ASCH, D.A. (2000). Pricing term insurance in the presence of a family history of breast or ovarian cancer. *North American Actuarial Journal*, **4**, 75–87.

LIEDE, A., MALIK, I.A., AZIZ, Z., DE LOS RIOS, P., KWAN, E. & NAROD, S.A. (2002). Contribution of BRCA1 and BRCA2 mutations to breast and ovarian cancer in Pakistan. *American Journal of Human Genetics*, **71**, 595–606.

LU, L., MACDONALD, A.S. & WATERS, H. (2006). Sampling distributions of critical illness insurance premium rates: Breast and ovarian cancer. *GIRC Research*

*Report 06/7.*

LYNCH, H.T., SNYDER, C., LYNCH, J.F., KARATPRAKLI, P., TROWONOU, A., METCALFE, K., NAROD, S.A. & GONG, G. (2006). Patient responses to the disclosure of BRCA mutation tests in hereditary breast-ovarian cancer families. *Cancer Genetics and Cytogenetics*, **165**, 91–97.

MAAT-KIEVIT, A., VEGTER-VAN DER VLIS, M., ZOETEWELJ, M., VAN HAERINGEN, A. & ROOS, R. (2000). Paradox of a better test for Huntington's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, **69**, 579–583.

MACDONALD, A.S. (1997). How will improved forecasts of individual lifetimes affect underwriting?. *Philosophical Transactions of the Royal Society B*, **352**, 1067–1075.

MACDONALD, A.S. (1999). Modelling the impact of genetics on insurance. *North American Actuarial Journal*, **3**, 83–101.

MACDONALD, A.S. (2000), Human Genetics and Insurance Issues, in *Bio-ethics for the New Millenium*, ed. Torrance, I., St. Andrew Press

MACDONALD, A.S., WATERS, H.R. & WEKWETE, C.T. (2003a). The genetics of breast and ovarian cancer I: A model of family history. *Scandinavian Actuarial Journal*, **2003**, 1–27.

MACDONALD, A.S., WATERS, H.R. & WEKWETE, C.T. (2003b). The genetics of breast and ovarian cancer II: A model of critical illness insurance. *Scandinavian Actuarial Journal*, **2003**, 28–50.

MACDONALD, A.S. & McIVOR, K. (2006). Application of a polygenic model of breast and ovarian cancer to critical illness insurance applications. *GIRC Research Report 06/2.*

MAGISTRONI, R., HE, N., WANG, K., ANDREW, R., JOHNSON, A., GABOW, P., DICKS, E., PARFREY, P., TORRA, R., SAN MILLAN, J.L., COTO, E., VAN DIJK, M., BREUNING, M., PETERS, D., BOGDANOVA, N., LIGABUE, G., ALBERTAZZI, A., HATEBOER, N., DEMETRIOU, K., PIERIDES, A., DELTAS, C., ST. GEORGE-HYSLOP, P., RAVINE, D. & PEI, Y. (2003). Genotype-renal func-

tion correlation in type 2 autosomal dominant polycystic kidney disease. *Journal of the American Society of Nephrology*, **14**, 1164–1174.

MALONE, K.E., DALING, J.R., THOMPSON, J.D., O'BRIEN, C.A., FRANCISCO, L.V. & OSTRANDER, E.A. (1998). Brca1 mutations and breast cancer in the general population: Analyses in women before age 35 years and in women before age 45 years with first-degree family history. *Journal of the American Medical Association*, **279**, 922–929.

MALONE, K.E., DALING, J.R., NEAL, C., SUTER, N.M., O'BRIEN, C., CUSHING-HAUGHEN, K., JONASDOTTIR, T.J., THOMPSON, J.D. & OSTRANDER, E.A. (2000). Frequency of BRCA1/BRCA2 mutations in a population-based sample of young breast carcinoma cases. *Cancer*, **88**, 1393–1402.

MARDER, K., ZHAO, H., MYERS, R., CUDKOWICZ, M., KAYSON, E., KIEBURTZ, K., ORME, C., PAULSEN, J., PENNEY, J., SIEMERS, E., SHOULSON, I. & THE HUNTINGTON STUDY GROUP (2000). Rate of functional decline in Huntington's disease. *Neurology*, **69**, 574–578.

MARRONI, F., ARETINI, P., D'ANDREA, E., CALIGO, M.A., CORTESI, L., VIEL, A., RICEVUTO, E., MONTAGNA, M., CIPOLLINI, G., FEDERICO, M., SANTAROSA, M., MARCHETTI, P., BAILEY-WILSON, J.E., BEVILACQUA, G., PARMIGIANI, G. & PRESCIUTTINI, S. (2004). Penetrances of breast and ovarian cancer in a large series of families tested for BRCA1/2 mutations. *European Journal of Human Genetics*, **12**, 899–906.

MASUDA, N., GOTO, J., MURAYAMA, N., WATANABE, M., KONDO, I. & KANAZAWA, I. (1995). Analysis of triplet repeats in the huntingtin gene in Japanese families affected with Huntington's disease. *Journal of Medical Genetics*, **32**, 701–705.

MCCORMICK, A.D., FLEMING, D. & CHARLTON, J. (1995). *Morbidity Statistics from General Practice: Fourth National Study 1991–1992*. Series MB5 No. 3. Washington, D.C.: OPCS, Government Statistical Service.

MCDONALD, S.P. & RUSS, G.R. (2002). Survival of recipients of cadaveric kidney transplants compared with those receiving dialysis treatment in Australia and New

Zealand, 1991 - 2001. *Nephrology Dialysis Transplantation*, **17**, 2212–2219.

MCNEIL, S.M., NOVELLETTO, A., SRINIDHI, J., BARNES, G., KORNBLUTH, I., ALTHERR, M.R., WASMUTH, J.J., GUSELLA, J.F., MACDONALD, M.E. & MYERS, R. (1997). Reduced penetrance of the Huntington's disease mutation. *Human Molecular Genetics*, **6**, 775–779.

MEEUWISSEN, P.A.M., SEYNAEVE, C., BREKELMANS, C.T.M., MEIJERS-HEIJBOER, H.J., KLIJN, J.G.M. & BURGER, C.W. (2005). Outcome of surveillance and prophylactic salpingo-oophorectomy in asymptomatic women at high risk for ovarian cancer. *Gynecologic Oncology*, **97**, 476–482.

MEIER-KRIESCHE, H-U., OJO, A.O., PORT, F.K., ARNDORFER, J.A., CIBRIK, D.M. & CAPLAN, B. (2001). Survival improvement among patients with end-stage renal disease: Trends over time for transplant recipients and wait-listed patients. *Journal of the American Society of Nephrology*, **12**, 1293–1296.

MENDEL, G. (1973), Experiments in plant-hybridization, in *Classic Papers in Genetics*, ed. Peters J.A., Englewood Cliffs, NJ: Prentice-Hall <sup>1</sup>

MIKI, Y., SWENSEN, J., SHATTUCK-EIDENS, D., FUTREAL, P.A., HARSHMAN, K., TAVTIGIAN, S., LIU, Q., COCHRAN, C., BENNETT, L.M., DING, W., BELL, R., ROSENTHAL, J., HUSSEY, C., TRAN, T., MCCLURE, M., FRYE, C., HATTIER, T., PHELPS, R., HAUGEN-STRANO, A., KATCHER, H., YAKUMO, K., GHOLAMI, Z., SHAFFER, D., STONE, S., BAYER, S., WRAY, C., BOGDEN, R., DAYANANTH, P., WARD, J., TONIN, P., NAROD, S., BRISTOW, P.K., NORRIS, F.H., HELVERING, L., MORRISON, P., ROSTECK, P., LAI, M., BARRETT, J.C., LEWIS, C., NEUHAUSEN, S., CANNON-ALBRIGHT, L., GOLDFAR, D., WISEMAN, R., KAMB, A. & SKOLNICH, M.H. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, **266**, 66–71.

MILUTINOVIC, J., FIALKOW, P.J., AGODOA, L.Y., PHILLIPS, L.A., RUDD,

---

<sup>1</sup>I have referenced this book as I it was easiest to find. The original reference for this paper is: MENDEL, G. (1865). Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines in Brünn*, **4**, 3–47., which appeared in 1866 and was translated into English by the Royal Horticultural Society (1901)

- T.G. & BRYANT, J.I. (1984). Autosomal dominant polycystic kidney disease: Symptoms and clinical findings. *Quarterly Journal of Medicine, New Series* **LIII**, 511–522.
- MITCHELL, R.J., CAMPBELL, H., FARRINGTON, S.M., BREWSTER, D.H., PORTEOUS, M.E. & DUNLOP, M.G. (2005). Prevalence of family history of colorectal cancer in the general population. *British Journal of Surgery*, **92**, 1161–1164.
- MOCHIZUKI, T., WU, G., HAYASHI, T., XENOPHONTOS, S.L., VELDHUISEN, B., SARIS, J.J., REYNOLDS, D.M., CAI, Y., GABOW, P.A., PIERIDES, A., KIMBERLING, W.J., BREUNING, M.H., DELTAS, C.C., PETERS, D.J. & SOMLO, S. (1996). PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. *Science*, **272**, 1339–1342.
- MORRISON, P., JOHNSTON, W. & NEVIN, N. (1995). The epidemiology of Huntington's disease in Northern Ireland. *Journal of Medical Genetics*, **32**, 524–530.
- MURPHY, S.W., FOLEY, R.N., BARRETT, B.J., KENT, G.M., MORGAN, J., BARRÉ, P., CAMPBELL, P., FINE, A., GOLDSTEIN, M.B., HANDA, S.P., JINDAL, K.K., LEVIN, A., MANDIN, H., MUIRHEAD, N., RICHARDSON, R.M.A. & PARFREY, P.S. (2000). Comparative mortality of hemodialysis and peritoneal dialysis in Canada. *Kidney International*, **57**, 1720–1726.
- MYERS, R., SAX, D.S., SCHOENFELD, M., BIRD, E., WOLF, P., VONSATTEL, J.P., WHITE, R.F. & MARTIN, J.B. (1985). Late onset of Huntington's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, **48**, 530–534.
- MYERS, R.H. (2004). Huntington's Disease genetics. *The Journal of the American Society of Experimental NeuroTherapeutics*, **1**, 255–262.
- NAROD, S.A., FORD, D., DEVILEE, P., BARKARDOTTIR, R.B., LYNCH, H.T., SMITH, S.A., PONDER, B.A.J., WEBER, B.L., GARBER, J.E., BIRCH, J.M., CORNELLIUS, R.S., KELSELL, D.P., SPURR, N.K., SMYTH, E., HAITES, N., SOBOL, H., BIGNON, Y.J., CHANG-CLAUDE, J., HAMANN, U., LINDBLOM, A., BORG, A., PIVER, M.S., GALLION, H.H., STRUEWING, J.P., WHITTEMORE, A., TONIN, P., GOLDGAR, D.E., EASTON, D.F. & THE BREAST CANCER LINKAGE CONSORTIUM (1995). An evaluation of genetic heterogeneity in 145 breast-

- ovarian cancer families. *American Journal of Human Genetics*, **56**, 254–264.
- NAROD, S.A. & FOULKES, W.D. (2004). BRCA1 and BRCA2: 1994 and beyond. *Nature Reviews*, **4**, 665–676.
- NELDER, J.A. & MEAD, R. (1965). A simplex method for function minimization. *The Computer Journal*, **7**, 308–313.
- NEUHAUSEN, S., GILEWSKI, T., NORTON, L., TRAN, T., MCGUIRE, P., SWENSEN, J., HAMPEL, H., BORGAN, P., BROWN, K., SKOLNICK, M., SHATTUCK EIDENS, D., JHANWAR, S., GOLDGAR, D. & OFFIT, K. (1996). Recurrent BRCA2 6174delT mutations in Ashkenazi Jewish women affected by breast cancer. *Nature Genetics*, **13**, 126–128.
- NEWCOMBE, R.G. (1981). A life table for onset of Huntington's chorea. *Annals of Human Genetics*, **45**, 375–85.
- NEWCOMBE, R.G., WALKER, D.A. & HARPER, P.S. (1981). Factors influencing age at onset and duration of survival in Huntington's chorea. *Annals of Human Genetics*, **45**, 387–96.
- NISHISHO, I., NAKAMURA, Y., MIYOSHI, Y., MIKI, Y., ANDO, H., HORII, A., KOYAMA, K., UTSUNOMIYA, J., BABA, S., HEDGE, P., MARKHAM, A., KRUSH, A.J., PETERSEN, G., HAMILTON, S.R., NILBERT, M.C., LEVY, D.B., BRYAN, T.M., PREISINGER, A.C., SMITH, K.J., SU, L-K., KINZLER, K.W. & VOGELSTEIN, B. (1991). Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science*, **253**, 665–669.
- ODDOUX, C., STRUEWING, J.P, CLAYTON, C.M., NEUHAUSEN, S., BRODY, L.C., KABACK, M., HAAS, B., NORTON, L., BORGAN, P., JHANWAR, S., GOLDGAR, D., OSTRER, H. & OFFIT, K. (1996). The carrier frequency of the BRCA2 6174delT mutation among Ashkenazi Jewish individuals is approximately 1%. *Nature Genetics*, **14**, 188–190.
- OFFIT, K., GILEWSKI, T., MCGUIRE, P., SCHLUGER, A., HAMPEL, H., BROWN, K., SWENSEN, J., NEUHAUSEN, S., SKOLNICK, M., NORTON, L. & GOLDGAR, D. (1996). Germline BRCA1 185delAG mutations in Jewish women

with breast cancer. *Lancet*, **347**, 1643–1645.

OFFICE FOR NATIONAL STATISTICS (O.N.S.) (1999). *Cancer 1971-1997 CD-ROM*. Office for National Statistics, London.

OFFICE OF POPOULATION CENSUSES AND SURVEYS (O.P.C.S.) (1991). *1990 Mortality Statistics: Cause Series DH 12 No 17*. London: O.P.C.S.

OFFICE OF POPOULATION CENSUSES AND SURVEYS (O.P.C.S.) (1993a). *1991 Mortality Statistics: Cause Series DH 12 No 18*. London: O.P.C.S.

OFFICE OF POPOULATION CENSUSES AND SURVEYS (O.P.C.S.) (1993b). *1992 Mortality Statistics: Cause Series DH 12 No 19*. London: O.P.C.S.

PANSE, F. (1942). *Die erbchorea; eine klinisch-genetisch studie*. Leipzig: Thieme.

PARFREY, P.S., BEAR, J.C., MORGAN, J., CRAMER, B.C., MCMANAMON, P.J., GAULT, M.H., CHURCHILL, D.N., SINGH, M., HEWITT, R., SOMLO, S. & REEDERS, S.T. (1990). The diagnosis and prognosis of autosomal dominant polycystic kidney disease. *The New England Journal of Medicine*, **323**, 1085–1090.

PASTERNAK, J. (1999). *An introduction to human molecular genetics*. Fitzgerald Science Press.

PENNEY, J.B., VONSATTEL, J.P., MACDONALD, M.E., GUSELLA, J.F. & MYERS, R.H. (1997). CAG repeat number governs the development rate of pathology in Huntington's disease. *Annals of Neurology*, **41**, 689–692.

PERRONE, R.D., RUTHAZER, R. & TERRIN, N.C. (2001). Survival after end-stage renal disease in autosomal dominant polycystic kidney disease: Contribution of extrarenal complications to mortality. *American Journal of Kidney Diseases*, **38**, 777–784.

PETERS, D.J.M., SPRUIT, L., SARIS, J.J., RAVINE, D., SANDKUIJL, L.A., FOSDAL, R., BOERSMA, J., VAN EIJK, R., NØRBY, S., CONSTANTINOU-DELTAS, C.D., PIERIDES, A., BRISSENDEN, J.E., FRANTS, R.R., VAN OMMEN, G.J.B. & BREUNING, M.H. (1993). Chromosome 4 localization of a second gene for autosomal dominant polycystic kidney disease. *Nature Genetics*, **5**, 359–362.



- PETERSEN, G.M., SLACK, J. & NAKAMURA, Y. (1991). Screening guidelines and premorbid diagnosis of familial adenomatous polyposis using linkage. *Gastroenterology*, **100**, 1658–1664.
- PETO, J., COLLINS, N., BARFOOT, R., SEAL, S., WARREN, W., RAHMAN, N., EASTON, D.F., EVANS, C., DEACON, J. & STRATTON, M.R. (1999). Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *Journal of the National Cancer Institute*, **91**, 943–949.
- PETO, J. (2002). Breast cancer susceptibility - a new look at an old model. *Cancer Cell*, **1**, 411–412.
- PETRUCELLI, N., DALY, M.B., BARS CULVER, J.O., LEVY-LAHAD, E. & FELDMAN, G.L. (Updated Dec 2005). BRCA1 and BRCA2 Hereditary Breast/Ovarian Cancer. In *Genereviews at Genetests: Medical Genetics Information Resource (database online)*. Copyright: University of Washington, Seattle. (1997–2006). Available at <http://www.genetests.org>.
- PIVER, M.S., BAKER, T.R., JISHI, M.F., SANDECKI, A.M., TSUKADA, Y., NATARAJAN, N., METTLIN, C.J. & BLAKE, C.A. (1993). Familial ovarian cancer: A report of 658 families from the Gilda Radner Familial Ovarian Cancer Registry 1981 - 1991. *Cancer*, **71**, 582–588.
- PONZ DE LEON, M., ANTONIOLO, A., ASCARI, A., ZANGHIERI, G. & SACCHETTI, G. (1987). Incidence and familial occurrence of colorectal cancer and polyps in a health care district of northern Italy. *Cancer*, **60**, 2848–2859.
- PORT, F.K., WOLFE, R.A., MAUGER, E.A., BERLING, D.P. & JIANG, K. (1993). Comparison of survival probabilities for dialysis patients vs cadaveric renal transplant recipients. *Journal of the American Medical Association*, **270**, 1339–1343.
- POWELL, S.M., PETERSEN, G.M., KRUSH, A.J., BOOKER, S., JEN, J., GIARDIELLO, F.M., HAMILTON, S.R., VOGELSTEIN, B. & KINZLER, K.W. (1993). Molecular diagnosis of familial adenomatous polyposis. *New England Journal of Medicine*, **329**, 1982–1987.



- PRESS, W.H., TEUKOLSKY, S.A., VETTERLING, W.T. & FLANNERY, B.P. (1992). *Numerical Recipes in C: The Scientific Art of Computing, 2nd Edition*. Cambridge University Press.
- RABBAT, C.G., THORPE, K.E., RUSSELL, J.D. & CHURCHILL, D.N. (2000). Comparison of mortality risk for dialysis patients and cadaveric first renal transplant recipients in Ontario, Canada. *Journal of the American Society of Nephrology*, **11**, 917–922.
- RAVINE, D., WALKER, R.G., GIBSON, R.N., FORREST, S.M., RICHARDS, R.I., FRIEND, K., SHEFFIELD, L.J., KINCAID-SMITH, P. & DANKS, D.M. (1992). Phenotype and genotype heterogeneity in autosomal dominant polycystic kidney disease. *The Lancet*, **340**, 1330–1333.
- REBBECK, T.R., LEVIN, A.M., EISEN, A., SNYDER, C., WATSON, P., CANNON-ALBRIGHT, L., ISAACS, C., OLOPADE, O., GARBER, J.E., GODWIN, A.K., DALY, M.B., NAROD, S.A., NEUHAUSEN, S.L., LYNCH, H.T. & WEBER, B.L. (1999). Breast cancer risk after bilateral prophylactic oophorectomy in BRCA1 mutation carriers. *Journal of the National Cancer Institute*, **91**, 1475–1479.
- REBBECK, T.R., LYNCH, H.T., NEUHAUSEN, S.L., NAROD, S.A., VAN’T VEER, L., GARBER, J.E., EVANS, G., ISAACS, C., DALY, M.B., MATLOFF, E., OLOPADE, O.I. & WEBER, B.L. FOR THE PREVENTION AND OBSERVATION OF SURGICAL END POINTS STUDY GROUP (2002). Prophylactic oophorectomy in carriers of BRCA1 or BRCA2 mutations. *The New England Journal of Medicine*, **346**, 1616–1621.
- REBBECK, T.R., FRIEBEL, T., LYNCH, H.T., NEUHAUSEN, S.L., VAN ’T VEER, L., GARBER, J.E., EVANS, G.R., NAROD, S.A., ISAACS, C., MATLOFF, E., DALY, M.B., OLOPADE, O.I. & WEBER, B.L. (2004). Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: The PROSE study group. *Journal of Clinical Oncology*, **22**, 1055–1062.
- REED, T.E., CHANDLER, J.H., HUGHES, E.M. & DAVIDSON, R.T (1958). Huntington’s chorea in Michigan. *American Journal of Human Genetics*, **10**, 201–225.
- REEDERS, S.T., BREUNING, M.H., DAVIES, K.E., NICHOLLS, R.D., JARMAN,

- A.P., HIGGS, D.R., PEARSON, P.L. & WEATHERALL, D.J. (1985). A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. *Nature*, **317**, 542–544.
- RISCH, H.A., MCLAUGHLIN, J.R., COLE, D.E.C., ROSEN, B., BRADLEY, L., KWAN, E., JACK, E., VESPRINI, D.J., KUPERSTEIN, G., ABRAHAMSON, J.L.A., FAN, I., WONG, B. & NAROD, S.A. (2001). Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. *American Journal of Human Genetics*, **68**, 700–710.
- ROA, B.B., BOYD, A.A., VOLCIK, K. & RICHARDS, C.S. (1996). Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nature Genetics*, **14**, 185–187.
- ROBINSON, B.H.B. AND HAWKINS, J.B. (1981). Procedures of the European Dialysis and Transplant Association. *17*, 20.
- ROMÃO, E.A., MOYSÉS NETO, M., TEIXEIRA, S.R., MUGLIA, V.F., VIEIRA-NETO, O.M. & DANTAS, M. (2006). Renal and extrarenal manifestations of autosomal dominant polycystic kidney disease. *Brazilian Journal of Medical and Biological Research*, **39**, 533–538.
- ROMEO, G., DEVOTO, M., COSTA, G., RONCUZZI, L., CATIZONE, L., ZUCHELLI, P., GERMINO, G.G., KEITH, T., WEATHERALL, D.J. & REEDERS, S.T. (1988). A second genetic locus for autosomal dominant polycystic kidney disease. *The Lancet*, **ii**, 8–10.
- ROOS, R.A.C., VAN DER VLIS, M.V., HERMANS, J., ELSHOVE, H.M., MOLL, A.C., VAN DE KAMP, J.J.P. & BRUYN, G.W. (1991). Age at onset in Huntington's disease: effect of line of inheritance and patient's sex. *Journal of Medical Genetics*, **28**, 515–519.
- ROOS, R.A.C., HERMANS, J., VAN DER VLIS, M.V., VAN OMMEN, G.J. & BRUYN, G.W. (1993). Duration of illness in Huntington's Disease is not related to age at onset. *Journal of Neurology, Neurosurgery and Psychiatry*, **56**, 98–100.
- ROSCOE, J.M., BRISSENDEN, J.E., WILLIAMS, E.A., CHERY, A.L. & SIL-

VERMAN, M. (1993). Autosomal dominant kidney disease in Toronto. *Kidney International*, **44**, 1101–1108.

ROSSETTI, S., STRMECKI, L., GAMBLE, V., BURTON, S., SNEDDON, V., PERAL, B., ROY, S., BAKKALOGLU, A., KOMEL, R., WINEARLS, C.G. & HARRIS, P.C. (2001). Mutation analysis of the entire PKD1 gene: genetic and diagnostic implications. *American Journal of Human Genetics*, **68**, 46–63.

ROSSETTI, S., CHAUVEAU, D., WALKER, D., SAGGAR-MALIK, A., WINEARLS, C.G., TORRES, V.E. & HARRIS, P.C. (2002). A complete mutation screen of the ADPKD genes by DHPLC. *Kidney International*, **61**, 1588–1599.

RUBINSZTEIN, D.C., LEGGO, J., COLES, R., ALMQVIST, E., BIANCALANA, V., CASSIMAN, J.J., CHOTAI, K., CONNARTY, M., CRAUFURD, D., CURTIS, A., CURTIS, D., DAVIDSON, M.J., DIFFER, A.M., DODE, C., DODGE, A., FRONTALI, M., RANEN, N.G., STINE, O.C., SHERR, M., ABBOTT, M.H., FRANZ, M.L., GRAHAM, C.A., HARPER, P.S., HEDREEN, J.C., JACKSON, A., KAPLAN, J.C., LOSEKOOT, M., MACMILLAN, J.C., MORRISON, P., TROTTER, Y., NOVELLETTO, A., SIMPSON, S.A., THEILMANN, J., WHITTAKER, J.L., FOLSTEIN, S.E., ROSS, C.A. & HAYDEN, M.R. (1996). Phenotypic characterization of individuals with 30-40 repeats in the Huntington's disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36-39 repeats. *American Journal of Human Genetics*, **59**, 16–22.

SCHRAG, D., KUNTZ, K.M., GARBER, J.E. & WEEKS, J.C. (1997). Decision analysis - effects of prophylactic mastectomy and oophorectomy on life expectancy among women with BRCA1 or BRCA2 mutations. *The New England Journal of Medicine*, **336**, 1465–1471.

SIMPSON, S.A., DAVIDSON, M.J. & BARRON, L.H. (1993). Huntington's disease in Grampian region: correlation of the CAG repeat number and the age of onset of the disease. *Journal of Medical Genetics*, **30**, 1014–1017.

SLATTERY, M.L. & KERBER, R.A. (1993). A comprehensive evaluation of family history and breast cancer risk. *Journal of the American Medical Association*, **270**, 1563–1568.

- SMITH, C. (1998). *Huntington's chorea: A mathematical model for life insurance*. Unpublished manuscript, Swiss Re, Zurich.
- SNELL, R.G., MACMILLAN, J.C., CHEADLE, J.P., FENTON, I., LAZAROU, L.P., DAVIES, P., MACDONALD, M.E., GUSELLA, J.F., HARPER, P.S. & SHAW, D.J. (1993). Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nature*, **4**, 393–397.
- STEVENS, D. (1977). *Huntington's chorea: A demographic genetic and clinical study*. MD Thesis, University of London.
- STEWART, J., DUNDAS, R., HOWARD, R., RUDD, A. & WOLFE, C. (1999). Ethnic differences in incidences of stroke: Prospective study with stroke register. *British Medical Journal*, **318**, 967–971.
- STINE, O.C., PLEASANT, N., FRANZ, M.L., ABBOTT, M.H., FOLSTEIN, S.E. & ROSS, C.A. (1993). Correlation between the onset age of Huntington's disease and length of the trinucleotide repeat in IT-15. *Human Molecular Genetics*, **2**, 1547–1549.
- ST JOHN, J.B., MCDERMOTT, F.T., HOPPER, J.L., DEBNEY, E.A., JOHNSON, W.R. & HUGHES, E.S.R. (1993). Cancer risk in relatives of patients with common colorectal cancer. *Annals of Internal Medicine*, **118**, 785–790.
- STRACHAN, T. & READ, A.P. (1999). *Human Molecular Genetics 2*. BIOS Scientific Publishers.
- STRATTON, M.R., FORD, D., NEUHAUSEN, S., SEAL, S., WOOSTER, R., FRIEDMAN, L.S., KING, M.C., EGILSSON, V., DEVILEE, P., MCMANUS, R., DALY, P.A., SMYTH, E., PONDER, B.A.J., PETO, J., CANNON-ALBRIGHT, L., EASTON, D.F. & GOLDGAR, D.E. (1994). Familial male breast cancer is not linked to the BRCA1 locus on chromosome 17q. *Nature Genetics*, **7**, 103–107.
- STRUEWING, J.P., BRODY, L.C., ERDOS, M.R., KASE, R.G., GIAMBARRESI, T.R., SMITH, S.A., COLLINS, F.S. & TUCKER, M.A. (1995). Detection of eight BRCA1 mutations in 10 breast/ovarian cancer families, including 1 family with male breast cancer. *American Journal of Human Genetics*, **57**, 1–7.

- STRUEWING, J.P., HARTGE, P., WACHOLDER, S., BAKER, S.M., BERLIN, M., MCADAMS, M., TIMMERMAN, M.M., BRODY, L.C. & TUCKER, M.A. (1997). The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *The New England Journal of Medicine*, **336**, 1401–1408.
- SUBRAMANIAN, K., LEMAIRE, J., HERSHEY, J.C., PAULY, M.V., ARMSTRONG, K. & ASCH, D.A. (2000). Estimating adverse selection costs from genetic testing for breast and ovarian cancer: The case of life insurance. *Journal of Risk and Insurance*, **66**, 531–550.
- SUJANSKY, E., BEELER KREUTZER, S., JOHNSON, A.M., LEZOTTE, D.C., SCHRIER, R.W. & GABOW, P.A. (1990). Attitudes of at-risk and affected individuals regarding presymptomatic testing for autosomal dominant kidney disease. *American Journal of Medical Genetics*, **35**, 510–515.
- SUTCLIFFE, S., PHAROAH, P.D.P., EASTON, D.F., PONDER, B.A.J. & THE UKCCR FAMILIAL OVARIAN CANCER STUDY GROUP (2000). Ovarian and breast cancer risks to women in families with two or more cases of ovarian cancer. *International Journal of Cancer*, **87**, 110–117.
- TAYLOR, S.D. (2004). Predictive genetic test decisions for Huntington's Disease: context, appraisal and new moral imperatives. *Social Science and Medicine*, **58**, 137–149.
- THOMPSON, E.A. (1993). Sampling and ascertainment in genetic epidemiology; a tutorial review. *Technical report No. 243. Department of Statistics, University of Washington*.
- TIRET, A., TAIEL-SARTRAL, M., TIRET, E. & LAROCHE, L. (1997). Diagnostic value of fundus examination in the familial adenomatous polyposis. *British Journal of Ophthalmology*, **81**, 755–758.
- TONIN, P., SEROVA, O., LENOIR, G., LYNCH, H., DUROCHER, F., SIMARD, J., MORGAN, K., NAROD, S. (1995). BRCA1 mutations in Ashkenazi Jewish women. *American Journal of Human Genetics*, **57**, 189.
- TORRA, R., BADENAS, C., DARNELL, A., NICOLAU, C., VOLPINI, V., REVERT,

- L. & ESTIVILL, X. (1996). Linkage, clinical features, and prognosis of autosomal dominant polycystic kidney disease types 1 and 2. *Journal of the American Society of Nephrology*, **7**, 2142–2151.
- TYLER, A., BALL, D. & CRAUFURD, D. (1992). Presymptomatic testing for Huntington's disease in the UK. *British Medical Journal*, **304**, 1593–1596.
- THE U.S.- VENEZUELA COLLABORATIVE RESEARCH PROJECT & WEXLER, N.S. (2004). Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 3498–3503.
- U.S. RENAL DISEASES SYSTEM (USRDS) (1999). Annual Report of the United States Renal Diseases System. *Washington, DC.: USRDS*.
- VASEN, H.F.A., GRIFFIOEN, G., OFFERHAUS, G.J.A., DEN HARTOG JAGER, F.C.A., VAN LEEUWEN-CORNELISSE, I.S.J., MEERA KHAN, P., LAMERS, C.B.H.W. & VAN SLOOTEN, E.A. (1990). The value of screening and central registration of families with familial adenomatous polyposis. *Diseases of the Colon and Rectum*, **33**, 227–231.
- VASEN, H.F., TESFAY, E., BOONSTRA, H., MOURITS, M.J., RUTGERS, E., VERHEYEN, R., OOSTERWIJK, J. & BEECH, L. (2005). Early detection of breast and ovarian cancer in families with BRCA mutations. *European Journal of Cancer*, **41**, 549–554.
- VENTERS, G. (1971). *Genetic and social aspects of Huntington's chorea*. PhD Thesis, University of Edinburgh.
- WALKER, D.A., HARPER, P.S., WELLS, C.E., TYLER, A., DAVIES, K. & NEWCOMBE, R.G. (1981). Huntington's Chorea in South Wales. A genetic and epidemiological study. *Clinical Genetics*, **19**, 213–21.
- WALLACE, D.C. (1972). Huntington's chorea in Queensland. A not uncommon disease. *The Medical Journal of Australia*, **1**, 299–307.
- WATSON, J.D. & CRICK, F.H. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.

- WENDT, G.G. & DROHM, C. (1972). Die Huntingtonsche chorea. Eine populations gentische studie. *Advances in Human Genetics*, **IV**, 1–21.
- WENDT, G.G., LANDZETTEL, I & UNTERREINER, I. (1959). Erkrankungsalter bei der Huntingtonschen Chorea. *Acta Genetica (Basel)*, **9**, 18–32.
- WINAWER, S.J., ZAUBER, A.G., GERDES, H., O'BRIEN, M.J., GOTTLIEB, L.S., STERNBERG, S.S., BOND, J.H., WAYE, J.D., SCHAPIRO, M., PANISH, J.F., KURTZ, R.C., SHIKE, M., ACKROYC, F.W., STEWART, E.T., SKOLNICK, M., BISHOP, D.T. & THE NATIONAL POLYP STUDY WORKGROUP (1996). Risk of colorectal cancer in the families of patients with adenomatous polyps. *New England Journal of Medicine*, **334**, 82–87.
- WOLFE, R.A., ASHBY, V.B., MILFORD, E.L., OJO, A.O., ETTENGER, R.E., AGODOA, L.Y.C., HELD, P.J. & PORT, F.K. (1999). Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *New England Journal of Medicine*, **341**, 1725–1730.
- WOOSTER, R., NEUHAUSEN, S.L., MANGION, J., QUIRK, Y., FORD, D., COLLINS, N., NGUYEN, K., SEAL, S., TRAN, T. & AVERILL, D., FIELDS, P., MARSHALL, G., NAROD, S., LENOIR, G.M., LYNCH, H., FEUNTEUN, J., DEVILEE, P., CORNELISSE, C.J., MENKO, F.H., DALY, P.A., ORMISTON, W., MCMANUS, R., PYE, C., LEWIS, C.M., CANNON-ALBRIGHT, L.A., PETO, J., PONDER, B.A.J., SKOLNICK, M.H., EASTON, D.F., GOLDGAR, D.E. & STRATTON, M.R. (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science*, **265**, 2088–2090.
- WOOSTER, R., BIGNELL, G., LANCASTER, J., SWIFT, S., SELA, S., MANGION, J., COLLINS, N., GREGORY, S., GUMBS, C., MICKLEM, G., BARFOOT, R., HAMOUDI, R., PATEL, S., RICE, C., BIGGS, P., HASHIM, Y., SMITH, A., CONNOR, F., ARASON, A., GUDMUNDSSON, J., FICENEC, D., KELSELL, D., FORD, D., TONIN, P., BISHOP, D.T., SPURR, N.K., PONDER, B.A.J., EELES, R., PETO, J., DEVILEE, P., CORNELISSE, C., LYNCH, H., NAROD, S., LENOIR, G., EGILSSON, V., BARKADOTTIR, R.B., EASTON, D.F., BENTLEY, D.R., FUTREAL, P.A., ASHWORTH, A & STRATTON, M.R. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature*, **378**, 789–792.

WORKING PAPER 14. Methodology Underlying the 1999-2002 CMI Critical Illness Experience Investigation

WORKING PAPER 18. 1999-2002 Critical Illness Experience: Feedback on Working Paper 14 and Future Work

WRIGHT, A.F., TEAGUE, P.W., POUND, S.E., PIGNATELLI, P.M., MACNICOL, A.M., CAROTHERS, A.D., DE MEY, R.J., ALLAN, P.L. & WATSON, M.L. (1993). A study of genetic linkage heterogeneity in 35 adult-onset polycystic kidney disease families. *Human Genetics*, **90**, 569–571.

WRIGHT, G.D., HUGHES, A.E., LARKIN, K.A., DOHERTY, C.C. & NEVIN, N.C. (1993). Genetic linkage analysis, clinical features and prognosis of autosomal dominant polycystic kidney disease in Northern Ireland. *Quarterly Journal of Medicine*, **86**, 459–463.



# Bibliography

ABROMOWITZ, M. & STEGUN, I.A. (1965). *Handbook of Mathematical Functions*. Cover Publications, New York.

THE AMERICAN COLLEGE OF MEDICAL GENETICS/AMERICAN SOCIETY OF HUMAN GENETICS HUNTINGTON'S DISEASE GENETIC TESTING WORKING GROUP (1998). Laboratory Guidelines for Huntington's Disease genetic testing. *American Journal of Human Genetics*, **62**, 1243–1247.

ARMSTRONG, K., WEBER, B., FITZGERALD, G., HERSHEY, J.C., PAULY, M.V., LEMAIRE, J., SUBRAMANIAN, K. & ASCH, D.A. (2003). Life insurance and breast cancer risk assessment: Adverse selection, genetic testing decisions, and discrimination. *American Journal of Medical Genetics*, **120A**, 359–364.

BRACKENRIDGE, R.D.C. & ELDER, W.J. (1998). *Medical Selection of Life Risks, Fourth Edition*. Macmillan Reference.

BURT, R.W. (1992). Accurate identification of familial adenomatous polyposis coli using newly developed genetic markers. *Gastroenterology*, **102**, A347.

BURTON, P.R., PALMER, L.J., JACOBS, K., KEEN, K.J., OLSON, J.M. & ELSTON, R.C. (2001). Ascertainment adjustment: Where does it take us?. *American Journal of Human Genetics*, **67**, 1505–1514.

CLAUS, E.B., SCHILDKRAUT, J.M., THOMPSON, W.D. & RISCH, N.J. (1996). The genetic attributable risk of breast and ovarian cancer. *Cancer*, **77**, 2318–2324.

DAVIS, P.J. & RABINOWITZ, P. (1975). *Methods of Numerical Integration*. Academic Press.

- DIXON, L.C.W. (1972). *Nonlinear Optimisation*. English University Press.
- ELSTON, R.C. & STEWART, J. (1971). A general model for the genetic analysis of pedigree data. *Human Heredity*, **21**, 523–542.
- EUROPEAN POLYCYSTIC KIDNEY DISEASE CONSORTIUM (1994). The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. *Cell*, **77**, 881–894.
- FLEMING, T.R. & HARRINGTON, D.P. (1991). *Counting Processes and Survival Analysis*. Wiley.
- GERBER, H.U. (1995). *Life Insurance Mathematics*. Springer.
- GREIG, D.M. (1980). *Optimisation*. Longman.
- HAIGH, B., HUQ, M. & HAYDEN, M.R. (Updated Aug 2005). Huntington Disease. In *GeneReviews at GeneTests: Medical Genetics Information Resource (database online)*. Copyright: University of Washington, Seattle. (1997-2006). Available at <http://www.genetests.org>.
- HARRIS, P.C. & TORRES, V.E. (Updated June 2006). Autosomal Dominant Polycystic Kidney Disease. In *GeneReviews at GeneTests: Medical Genetics Information Resource (database online)*. Copyright: University of Washington, Seattle. (1997-2006). Available at <http://www.genetests.org>.
- HARPER, P.S. & NEWCOMBE, R.G. (1992). Age at onset and life table risks in genetic counselling for Huntington's disease.. *Journal of Medical Genetics*, **29**, 239–42.
- HARPER, P.S., TYLER, A., SMITH, S., JONES, P., NEWCOMBE, R.G. & MCBROOM, V. (1982). A genetic register for Huntington's chorea in South Wales.. *Journal of Medical Genetics*, **19**, 241–5.
- HOEM, J.M. (1988). The versatility of the Markov chain as a tool in the mathematics of life insurance. *Transactions of the 23rd International Congress of Actuaries, Helsinki* **5**, 171–202.
- JACOBSEN, M. (1982). *Statistical Analysis of Counting Processes*. Springer.

- JAMES, C.M., HOULIHAN, G.D., SNELL, R.G., CHEADLE, J.P. & HARPER, P.S. (1994). Late-onset Huntington's disease: A clinical and molecular study. *Age Ageing*, **23**, 445–8.
- LE GRYS, D.J. (1997). Actuarial considerations on genetic testing. *Philosophical Transaction of the Royal Society B*, **352**, 1057–1061.
- MACDONALD, A.S. & PRITCHARD, D.J. (2000). A mathematical model of Alzheimer's disease and the APOE gene. *ASTIN Bulletin*, **30**, 69–110.
- MACDONALD, A.S. & PRITCHARD, D.J. (2001). Genetics, Alzheimer's disease and long-term care insurance. *North American Actuarial Journal*, **5:2**, 54–78.
- MACDONALD, A.S. (2002). Genetics and health costs: some actuarial models. *Law, Probability and Risk*, **1**, 97–118.
- MACDONALD, A.S. (2003). Moratoria on the Use of Genetic Tests and Family History for Mortgage-Related Life Insurance. *British Actuarial Journal*, **9**, 217–237.
- MACDONALD, A.S. (2003). Genetics and Insurance: What Have We Learned So Far?. *Scandinavian Actuarial Journal*, **2003**, 324–348.
- MACDONALD, A.S. & YANG, S.Y. (2003). Highly Penetrant Genetic Disorders: An Atlas for Critical Illness Insurance. *British Actuarial Journal*, **9**, 875–901. Additional information (Occasional Actuarial Research Document No. 42, deposited in the libraries of the Faculty and Institute of Actuaries).
- MCCONNELL, C.R. (1990). Pocket computer approximation for areas under the standard Normal curve. *American Statistician*, **44**, 63.
- PETERSEN, G.M. (1994). Knowledge of the adenomatous polyposis coli gene and its clinical application. *Annals of Medicine*, **26**, 205–208.
- POTTER, N.T., SPECTOR, E.B. & PRIOR, T.W. (2004). Technical standards and guidelines for Huntington disease testing. *Genetics in Medicine*, **6**, 61–65.
- QUARRELL, O.W.J., TYLER, A., JONES, M.P., NORDIN, M. & HARPER, P.S. (1988). Population studies of Huntington's disease in Wales. *Clinical Genetics*, **33**,

189–195.

SOLOMON, C. & BURT, R.W. (Updated Oct 2005). APC-Associated Polyposis Conditions. In *GeneReviews at GeneTests: Medical Genetics Information Resource (database online)*. Copyright: University of Washington, Seattle. (1997-2006). Available at <http://www.genetests.org>.

STEIN, L.D. (2004). Human genome: End of the beginning. *Nature*, **431**, 915 – 916.

STRACHAN, T. (1992). *The Human Genome*. BIOS Scientific Publishers.

DE VYLDER, F.E. (1997). *Life Insurance Theory: Actuarial Perspectives*. Kluwer Academic Publishers.

DE VYLDER, F.E. GOOVAERTS, M. & HAEZENDONCK, J. (1983). *Premium Calculation in Insurance*. D. Reidel Publishing Company.

YOUNG, A.B. (2003). Huntingtin in health and disease. *Journal of Clinical Investigation*, **111**, 299–302.